

# An Upgrading Algorithm with Optimal Power Law

Or Ordentlich, *Member, IEEE*, Ido Tal, *Senior Member, IEEE*

**Abstract**—Consider a channel  $W$  along with a given input distribution  $P_X$ . In certain settings, such as in the construction of polar codes, the output alphabet of  $W$  is ‘too large’, and hence we replace  $W$  by a channel  $Q$  having a smaller output alphabet. We say that  $Q$  is upgraded with respect to  $W$  if  $W$  is obtained from  $Q$  by processing its output. In this case, the mutual information  $I(P_X, W)$  between the input and output of  $W$  is upper-bounded by the mutual information  $I(P_X, Q)$  between the input and output of  $Q$ . In this paper, we present an algorithm that produces an upgraded channel  $Q$  from  $W$ , as a function of  $P_X$  and the required output alphabet size of  $Q$ , denoted  $L$ . We show that the difference in mutual informations is ‘small’. Namely, it is  $O(L^{-2/(|\mathcal{X}|-1)})$ , where  $|\mathcal{X}|$  is the size of the input alphabet. This power law of  $L$  is optimal. We complement our analysis with numerical experiments which show that the developed algorithm improves upon the existing state-of-the-art algorithms also in non-asymptotic setups.

## I. INTRODUCTION

In his seminal paper on polar codes, Arikan introduced synthetic channels [2, equation (5)], also called bit-channels. These synthetic channels have a binary input alphabet and an intractably large output alphabet. Namely, the output alphabet size of such a channel is at least  $2^N$ , where  $N$  is the length of the polar code. When decoding a polar code, this large set does not pose a problem: the decoder must contend with only a single element from the output alphabet. Neither is the output alphabet size a problem when encoding a polar code. However, when *constructing* a polar code, the vast size of the output alphabet is very much an issue. Since polar codes have since been generalized in many ways, let us call the polar codes introduced in [2] ‘vanilla polar codes’. To construct a vanilla polar code, one has to pick the ‘almost noiseless’ synthetic channels. That is, essentially, to calculate the probability of a maximum-likelihood (ML) decoder misdecoding the input to the synthetic channel, upon seeing the output of the channel. Clearly, the trivial method of iterating over all possible outputs in order to calculate this quantity will not work, since we cannot iterate over an intractably large set.

A crucial observation is that instead of considering the original synthetic channel, one may approximate it by another channel having a much smaller output alphabet size [3]. Namely, if a channel has a manageable output alphabet size, we may directly calculate the probability of ML misdecoding. A further observation is that if the approximating channel

is degraded with respect to the synthetic channel — the approximating channel can be obtained by processing the output of the synthetic channel — then the probability of misdecoding in the synthetic channel is upper bounded by the probability of misdecoding in the approximating channel.

These observations, combined with [4, Lemma 1.8], were used in [5] for constructing vanilla polar codes. The key part was an algorithm which transformed a given channel into a degraded approximating channel with a prescribed output alphabet size. The approximating channel was attained by quantizing the output of the synthetic channel to a relatively small alphabet, while decreasing the mutual information between the input and quantized output as little as possible, compared to the mutual information between the input and un-quantized output. This algorithm was used, successively, to approximate each of the  $N$  synthetic channels by a corresponding degrading channel. The  $k$  approximating channels with the smallest probability of ML misdecoding were used to construct a vanilla polar code with a prescribed dimension  $k$ . Summing the ML misdecoding probabilities of these  $k$  approximating channels gives an upper bound on the probability of misdecoding the polar code using the successive cancellation decoder.

If  $W$  is degraded with respect to  $Q$ , we will also say that  $Q$  is upgraded with respect to  $W$ . In [5], a companion algorithm, approximating a given channel by an upgraded channel with a prescribed output alphabet size was also given. In the context of the vanilla polar codes presented in [2], the importance of this companion algorithm was rather secondary. That is, it resulted in a lower bound on the misdecoding probability of the synthetic channel, and thus one could gauge, through the sandwich property, the closeness of the approximation of the misdecoding probability.

Shortly after their introduction in [2], polar codes were generalized in various directions. Three important generalizations that soon followed are those to lossy compression [6], asymmetric channels [7], and the wiretap channel [8]–[12]. In these settings, the construction of the polar code also calls for searching for synthetic channels that are ‘very noisy’. That is, a maximum a-posteriori (MAP) decoder trying to guess the input to the channel given the output must have a probability of failure close to  $1/2$ . The same problem of an intractably large output alphabet manifests itself, and the same solution of approximating the channel may be called upon, save for one difference: we now use the upgrading algorithm in [2] in order to lower bound the above probability of MAP misdecoding. Thus, in these contexts, an upgrading approximation is a key part of constructing the polar code, and ceases to be a tool of secondary importance. A more recent example of a polar coding variant in which upgrading plays a key role in the construction process is that of polar codes for settings with memory [13]–[17].

Or Ordentlich is with the School of Computer Science and Engineering, Hebrew University of Jerusalem, Israel (email: or.ordentlich@mail.huji.ac.il). Ido Tal is with the Viterbi Department of Electrical Engineering, Technion, Haifa 32000, Israel (email: idotal@ee.technion.ac.il).

The work of Or Ordentlich was supported by the ISF under Grant 1791/17. The work of Ido Tal was supported by the BSF under Grant 2018218. The material in this paper was presented in part at the 2020 International Zurich Seminar on Information and Communication [1].

A partial list of papers relating to upgrading and degrading approximations is [18]–[31]. Specifically, [27] contains a generalization of the upgrading approximation presented in [5] to cases in which the input has a non-uniform binary distribution.

The main contribution of this paper is the development and analysis of a novel upgrading algorithm for a general input distribution on an alphabet with cardinality greater than two. Our key idea is to apply a reduction to the binary case, inspired by [28], in order to use the algorithm in [27]. Our algorithm will have an optimal power law, a concept we will shortly define.

As a secondary contribution, we show that the channel degrading/quantization algorithm proposed in [28, Proof of Proposition 7] for the non-binary input case, has computational complexity which is significantly smaller than the previous state-of-the-art degrading algorithm. Along the way, we also improve the constants in the best known bounds on the degrading and upgrading costs for binary input alphabets under the greedy merge and greedy split algorithms.

The paper is organized as follows: In Section II we recall the definition of the channel upgrading problem, while in Section III we reformulate it as an optimization problem on Markov chains satisfying several requirements. Our algorithm, along with its performance guarantees is presented in Section IV, and the analysis is carried out in Section V. In Section VI we turn our attention to the channel degrading/quantization problem and describe a degrading algorithm for a non-binary input alphabet, that attains the optimal power law with complexity that is near-linear in the channel's output cardinality. Finally, In Section VII we demonstrate the practical benefits of our upgrading, as well as degrading, algorithms through numerical experiments. In the analysis of the upgrading and degrading costs of the algorithms we present, we use upper bounds on the upgrading and degrading cost for the case of a binary input. Such bounds with optimal power laws have already appeared in [27]. In the Appendix we bring somewhat simplified derivations of those bounds, that further result in improved constants.

## II. SETTING

We are given a channel  $W : \mathcal{X} \rightarrow \mathcal{Y}$ , where  $\mathcal{X}$  is termed the input alphabet and  $\mathcal{Y}$  is termed the output alphabet. We denote the probability of receiving  $y \in \mathcal{Y}$  given that  $x \in \mathcal{X}$  was transmitted by  $W(y|x)$ . We are also given an input distribution  $P_X$ . That is, we denote by  $P_X(x)$  the probability that  $x \in \mathcal{X}$  was the input to the channel. In this paper, we will assume that both  $\mathcal{X}$  and  $\mathcal{Y}$  are finite.<sup>1</sup> We denote the mutual information between the input and output of  $W$  as

$$I(P_X, W) \triangleq I(X; Y),$$

where  $X$  and  $Y$  are random variables with joint distribution

$$P_{X,Y}(x, y) = P_X(x)W(y|x). \quad (1)$$

Let  $Q : \mathcal{X} \rightarrow \mathcal{Z}$  be a channel with the same input alphabet as  $W : \mathcal{X} \rightarrow \mathcal{Y}$ . We say that  $Q$  is *upgraded* with respect to  $W$  if we can obtain  $W$  by processing the output of  $Q$ . That

is, if there exists a third channel  $\Phi : \mathcal{Z} \rightarrow \mathcal{Y}$  such that, for every  $x \in \mathcal{X}$  and  $y \in \mathcal{Y}$ ,

$$W(y|x) = \sum_{z \in \mathcal{Z}} Q(z|x)\Phi(y|z).$$

Our goal in this paper is, given an input distribution  $P_X$ , a channel  $W : \mathcal{X} \rightarrow \mathcal{Y}$ , and a parameter  $L$ , to construct a channel  $Q : \mathcal{X} \rightarrow \mathcal{Z}$  that is upgraded with respect to  $W$  and whose output alphabet size satisfies  $|\mathcal{Z}| \leq L$ . Many such channels  $Q$  exist. By the data processing inequality, it must hold that

$$I(P_X, Q) \geq I(P_X, W).$$

Hence, our figure of merit of how well  $Q$  approximates  $W$  will be the difference  $I(P_X, Q) - I(P_X, W)$ , which we wish to minimize. It turns out that there exist pairs of input distributions and channels for which such an approximation is inherently ‘hard’. That is, [27, Section IV] shows a  $P_X$  and  $W$  for which<sup>2</sup>

$$I(P_X, Q) - I(P_X, W) = \Omega(L^{-2/(|\mathcal{X}|-1)}), \quad (2)$$

for every valid choice of  $Q$ . Our method will produce, for any  $P_X$  and  $W$ , a  $Q$  for which

$$I(P_X, Q) - I(P_X, W) = O(L^{-2/(|\mathcal{X}|-1)}). \quad (3)$$

Hence, (2) and (3) imply that our algorithm has an optimal power law. We note that in this paper,  $|\mathcal{X}|$  is assumed to be a fixed constant. That is, in the asymptotic notation used in (2) and (3), the multiplying constant generally *does* depend on  $|\mathcal{X}|$ .

Two comments are in order. First, recall that we have assumed that the output alphabet of  $W$  is finite. In many important settings, this will not be the case. For example, if  $W$  models the addition of Gaussian noise to a given input supported on a finite alphabet. In such a case, we may use the upgrading algorithm in [24], as a preliminary step. That is, the algorithm in [24] constructs an upgraded channel  $Q$  satisfying  $I(P_X, Q) - I(P_X, W) = O(L^{-1/(|\mathcal{X}|-1)})$ . Note that this expression is worse than (3), since the  $-2$  in (3) has been replaced by  $-1$ . However, for the preliminary step, we may run [24] with  $L^2$  in place of  $L$ , and then run the algorithm we will shortly introduce on the resulting channel. Clearly, the difference between the original  $W$  and the final  $Q$  will be  $O(L^{-2/(|\mathcal{X}|-1)})$ . Also, it is easy to see that since we have applied two upgrading operations in series, the final channel is upgraded with respect to the initial one. Second, we would like to stress that our algorithm does not generally find the  $Q$  which minimizes  $I(P_X, Q) - I(P_X, W)$ . The problem of finding such a  $Q$  has been solved for  $|\mathcal{X}| = 2$  in [27, Section V], but remains open for  $|\mathcal{X}| > 2$ .

## III. MARKOV CHAIN REPRESENTATION

### A. Distributions notation

In this paper, several probability distributions will be defined and used. For example, we will denote by  $P_{X,Z,Y}^*(x, z, y)$

<sup>2</sup>For sequences  $a = a(L)$  and  $b = b(L)$ , we write  $a(L) = O(b(L))$  if there exists constants  $C > 0$  and  $L_0$  such that  $a(L) \leq Cb(L)$  for all  $L \geq L_0$ . Similarly,  $a(L) = \Omega(b(L))$  means  $a(L) \geq Cb(L)$ .

<sup>1</sup>But see the discussion below about infinite sized alphabets.

a probability distribution on  $x \in \mathcal{X}$ ,  $z \in \mathcal{Z}$ , and  $y \in \mathcal{Y}$ . The subscript  $X, Z, Y$  serves two purposes. The first purpose is to detail how the ‘corresponding random variables’ are defined. Namely, since the first, second, and third entries in the subscript (function arguments list) are  $X(x)$ ,  $Z(z)$ ,  $Y(y)$ , the corresponding random variables are  $X$ ,  $Z$ , and  $Y$ . Also, the probability of  $X = x$ ,  $Z = z$  and  $Y = y$  equals  $P_{X,Z,Y}^*(x, z, y)$ . We denote the probability of this event as

$$\mathbb{P}(X = x, Z = z, Y = y) = P_{X,Z,Y}^*(x, z, y).$$

We will always detail under which probability distribution  $\mathbb{P}$  is calculated. The second purpose of the subscript  $X, Z, Y$  is to define derived probability distributions. For example,  $P_{Z|Y}^*(z|y)$  denotes the probability that  $Z = z$  given that  $Y = y$ , where  $Z$  and  $Y$  are the random variables described earlier. That is,

$$\begin{aligned} P_{Z|Y}^*(z|y) &= \mathbb{P}(Z = z|Y = y) \\ &= \frac{\sum_{x \in \mathcal{X}} P_{X,Z,Y}^*(x, z, y)}{\sum_{x \in \mathcal{X}, z' \in \mathcal{Z}} P_{X,Z,Y}^*(x, z', y)}. \end{aligned}$$

To avoid corner cases (such as division by zero in the above definition), we will always assume without loss of generality (w.l.o.g.) that there are no symbols with probability zero. That is, for each  $x \in \mathcal{X}$  it holds that  $P_X^*(x) > 0$  (otherwise we could remove  $x$  from  $\mathcal{X}$ ), etc.

### B. Restatement of setting

We find it conceptually simpler<sup>3</sup> to merge the input distribution  $P_X(x)$  and channel  $W(y|x)$  into one joint distribution  $P_{X,Y}(x, y)$  as in (1). We will call this the ‘given distribution’. We now restate our setting as follows.

- 1) We are given a distribution  $P_{X,Y}(x, y)$ , where  $x \in \mathcal{X}$  and  $y \in \mathcal{Y}$ , both sets being finite.
- 2) We must find a distribution  $P_{X,Z,Y}^*(x, z, y)$ , where  $x \in \mathcal{X}$ ,  $z \in \mathcal{Z}$ ,  $y \in \mathcal{Y}$ . We require that  $|\mathcal{Z}| \leq L$ .
- 3) The marginalization of  $P_{X,Z,Y}^*(x, z, y)$  over  $z$  must produce  $P_{X,Y}(x, y)$ . Namely, for all  $x \in \mathcal{X}$  and  $y \in \mathcal{Y}$ ,

$$P_{X,Y}(x, y) = \sum_{z \in \mathcal{Z}} P_{X,Z,Y}^*(x, z, y).$$

- 4) The corresponding random variables  $X, Z, Y$  must form a Markov chain. Namely, we may write, for all  $x \in \mathcal{X}$ ,  $z \in \mathcal{Z}$ ,  $y \in \mathcal{Y}$ ,

$$P_{X,Z,Y}^*(x, z, y) = P_X^*(x)P_{Z|X}^*(z|x)P_{Y|Z}^*(y|z). \quad (4)$$

That is, the last term  $P_{Y|Z}^*(y|z)$  is not a function of  $x$ ; our processing of  $Z$  to form  $Y$  is done without knowledge of  $X$ .

<sup>3</sup>Indeed, one might argue that this is more natural. That is, if we were to coerce the input-distribution/channel terminology to the construction of polar codes for settings with memory [14]–[16], using our algorithm for the construction of a polar code would entail defining the ‘channel input’ as consisting of the actual symbol that was fed to the channel, along with a pair of hidden states encapsulating the combined channel and input process state at the beginning and end of the relevant block. Namely, with respect to [15, equation (25)], the ‘channel input’ is  $(U_i, S_0, S_N)$  and the channel output is  $Q_i$ .

- 5) Our figure of merit is the difference

$$\begin{aligned} I(X; Z) - I(X; Y) &= H(X|Y) - H(X|Z) \\ &= I(X; Z|Y), \end{aligned} \quad (5)$$

where the second equality follows by the Markov property.

A simple observation that will greatly simplify our derivations is that  $X - Z - Y$  form a Markov chain in this order iff  $Y - Z - X$  do. That is, we may replace (4) by the equivalent condition

$$P_{X,Z,Y}^*(x, z, y) = P_Y^*(y)P_{Z|Y}^*(z|y)P_{X|Z}^*(x|z). \quad (6)$$

## IV. THE ALGORITHM

### A. Upgradation for the Binary Case

Our key idea is to apply a reduction from the case in which the input alphabet  $\mathcal{X}$  is non-binary to a case in which the input alphabet is binary. In aid of this, we recall that in [27, Section VI], an efficient upgrading algorithm, termed ‘greedy-split’ for the binary-input case is presented. Namely, denote by

$$\mathcal{X}' = \{0, 1\}$$

the binary alphabet. Then, given a joint distribution  $P_{X',Y}(x', y)$ , where  $x' \in \mathcal{X}'$  and  $y \in \mathcal{Y}$ , we have an algorithm which produces, for a given  $L$ , a distribution  $P_{X',Z',Y}^*(x', z', y)$  such that the conditions in Section III are fulfilled, and the mutual information difference is  $O(L^{-2})$ .

In Appendices A and B we recall the greedy-split algorithm from [27], and give a full self contained proof for the bound  $I(X'; Z') - I(X'; Y) = O(L^{-2})$ , with improved constants compared to [27, Theorem 17].<sup>4</sup> Namely, we prove the following.<sup>5</sup>

*Theorem 1:* Let  $(X, Y) \sim P_{X,Y}$  be random variables in  $\mathcal{X} \times \mathcal{Y}$ , where  $|\mathcal{X}| = 2$  and  $\mathcal{Y}$  is discrete. For any natural  $L \geq 2$ , there exists a random variable  $Z$  of cardinality  $|\mathcal{Z}| = L$ , such that  $X - Z - Y$  form a Markov chain in this order, and

$$I(X; Z) - I(X; Y) \leq 128L^{-2}. \quad (7)$$

Our proof follows that of [27, Theorem 17], essentially step by step, but is simpler and shorter. This is mainly due to a simplification of the ‘sphere-packing’ argument. In [27], it was argued that given  $n$  distributions  $P_1, \dots, P_n$  in the  $(q - 1)$ -dimensional simplex, if  $n$  is sufficiently large, we must be able to find  $P_i$  and  $P_j$ ,  $1 \leq i < j \leq n$ , such that  $H(\alpha P_i + (1 - \alpha)P_j) - \alpha H(P_i) - (1 - \alpha)H(P_j)$  is small, for all  $\alpha \in [0, 1]$ . The argument leading to this conclusion was of a sphere-packing nature. As the same argument was needed in [27] also for the analysis of a channel degradation algorithm, termed ‘greedy-merge’, for general alphabet sizes, the sphere-packing argument was derived for general  $q$ . This led to various technical complications, which in turn led to

<sup>4</sup>The bound proved in [27, Theorem 17] was  $I(X; Z) - I(X; Y) \leq 2\nu(2)L^{-2}$ , where  $\nu(2) = 64/(\sqrt{3}/2 - 1)^2$ . Thus, we improve the constant by a factor of  $(\sqrt{3}/2 - 1)^{-2} \approx 19.79$ .

<sup>5</sup>Throughout, all logarithms are taken to the natural base.

rather loose constants. Restricting attention to  $q = 2$ , the simplex reduces to the  $[0, 1]$  interval, and the derivation of the sphere-packing bound is significantly simplified, leading also to better constants in the bound.

### B. Upgradation for the General Case

Denote the input alphabet size as

$$|\mathcal{X}| = q.$$

Similarly to the method in [28], we will now use the ‘one-hot’ representation of  $x \in \mathcal{X}$  to affect the reduction. Namely, w.l.o.g. let us assume that

$$\mathcal{X} = \{1, 2, \dots, q\}.$$

We will replace  $x \in \mathcal{X}$  by a length  $q - 1$  vector  $g(x) = (x_1, x_2, \dots, x_{q-1})$ , such that

$$x_i = \begin{cases} 1 & \text{if } x = i \\ 0 & \text{otherwise} \end{cases}$$

Namely, for  $1 \leq i \leq q - 1$ , we map  $x = i$  to the vector  $g(x)$  of length  $q - 1$  that has entry  $i$  equal to 1, and all other entries equal to 0. We map  $x = q$  to  $g(q)$ , the all-zero vector of length  $q - 1$ . Since the mapping  $g$  is invertible, we will often abuse notation and simply write  $x = (x_1, x_2, \dots, x_{q-1})$ .

Given the joint distribution  $P_{X,Y}$ , let  $X$  and  $Y$  be corresponding random variables. Recalling our convention of denoting  $X = (X_1, X_2, \dots, X_{q-1})$ , our first step is to define the following  $q - 1$  joint distributions: for  $1 \leq i \leq q - 1$ , let

$$\begin{aligned} \alpha_{X_i, Y}^{(i)}(x', y) &\triangleq \mathbb{P}(X_i = x', Y = y | X_1^{i-1} = 0_1^{i-1}) \\ &= \mathbb{P}(Y = y | X_1^{i-1} = 0_1^{i-1}) \mathbb{P}(X_i = x' | Y = y, X_1^{i-1} = 0_1^{i-1}) \\ &= \alpha_Y^{(i)}(y) \alpha_{X_i | Y}^{(i)}(x' | y), \end{aligned} \quad (8)$$

where  $0_1^{i-1}$  is the all-zero vector of length  $i - 1$ , and

$$\begin{aligned} \alpha_Y^{(i)}(y) &\triangleq \mathbb{P}(Y = y | X_1^{i-1} = 0_1^{i-1}) \\ \alpha_{X_i | Y}^{(i)}(x' | y) &\triangleq \mathbb{P}(X_i = x' | Y = y, X_1^{i-1} = 0_1^{i-1}). \end{aligned}$$

Note that if  $i = 1$ , there is no conditioning. We apply the binary-input upgrading algorithm to each of the above distributions, but require that the resulting output alphabet size be at most

$$\Lambda = \left\lfloor L^{1/(q-1)} \right\rfloor. \quad (9)$$

Thus, for each  $1 \leq i \leq q - 1$ , the binary-upgrading algorithm returns a distribution

$$\beta_{X_i, Z_i, Y}^{(i)}(x', z', y),$$

where  $x' \in \mathcal{X}'$ ,  $z' \in \mathcal{Z}^{(i)}$ ,  $y \in \mathcal{Y}$ , and the size of  $\mathcal{Z}^{(i)}$  satisfies  $|\mathcal{Z}^{(i)}| \leq \Lambda$ . By definition of upgrading, the induced random variables satisfy the Markov chain  $Y - Z_i - X_i$  in this order, and we may therefore write this distribution as

$$\begin{aligned} \beta_{X_i, Z_i, Y}^{(i)}(x', z', y) &= \beta_Y^{(i)}(y) \beta_{Z_i | Y}^{(i)}(z' | y) \beta_{X_i | Z_i}^{(i)}(x' | z') \\ &= \alpha_Y^{(i)}(y) \beta_{Z_i | Y}^{(i)}(z' | y) \beta_{X_i | Z_i}^{(i)}(x' | z'), \end{aligned}$$

and furthermore, we have that the concatenation of the channels  $\beta_{Z_i | Y}^{(i)} : \mathcal{Y} \rightarrow \mathcal{Z}^{(i)}$  and  $\beta_{X_i | Z_i}^{(i)} : \mathcal{Z}^{(i)} \rightarrow \mathcal{X}'$  results in the

channel  $\alpha_{X_i | Y}^{(i)} : \mathcal{Y} \rightarrow \mathcal{X}'$ . Namely, for all  $y \in \mathcal{Y}$  and  $x' \in \mathcal{X}'$ , we have that

$$\alpha_{X_i | Y}^{(i)}(x' | y) = \sum_{z' \in \mathcal{Z}^{(i)}} \beta_{Z_i | Y}^{(i)}(z' | y) \beta_{X_i | Z_i}^{(i)}(x' | z'). \quad (10)$$

We also recall for future use that the corresponding random variables satisfy (7), with  $\Lambda$  in place of  $L$ .

We use the above  $q - 1$  distribution in order to define the distribution  $P_{X, Z, Y}^*$ . Denote  $z = (z_1, z_2, \dots, z_{q-1})$ . Also, recall our one-hot convention for  $x$ , namely  $x = (x_1, x_2, \dots, x_{q-1})$ . Then, for

$$\mathcal{Z} = \mathcal{Z}^{(1)} \times \mathcal{Z}^{(2)} \times \dots \times \mathcal{Z}^{(q-1)},$$

we define for  $x \in \mathcal{X}$ ,  $z \in \mathcal{Z}$ , and  $y \in \mathcal{Y}$ ,

$$\begin{aligned} P_{X, Z, Y}^*(x, z, y) &= P_Y(y) \cdot \left( \prod_{i=1}^{q-1} \beta_{Z_i | Y}^{(i)}(z_i | y) \right) \\ &\quad \cdot \left( \prod_{i=1}^{q-1} \gamma_{X_i | Z_i, X_1^{i-1}}^{(i)}(x_i | z_i, x_1^{i-1}) \right), \end{aligned} \quad (11)$$

where, for  $1 \leq i \leq q$ ,

$$\begin{aligned} \gamma_{X_i | Z_i, X_1^{i-1}}^{(i)}(x_i | z_i, x_1^{i-1}) &= \begin{cases} \beta_{X_i | Z_i}^{(i)}(x_i | z_i) & \text{if } x_1^{i-1} = 0_1^{i-1}, \\ 1 & \text{if } x_1^{i-1} \neq 0_1^{i-1} \text{ and } x_i = 0, \\ 0 & \text{otherwise.} \end{cases} \end{aligned} \quad (12)$$

To summarize, our algorithm consists of three steps:

- 1) Construct the distributions  $\alpha_{X_i, Y}^{(i)}(x', y)$  for  $i = 1, \dots, q - 1$ , according to (8). Each such distribution is supported on  $\{0, 1\} \times \mathcal{Y}$ ;
- 2) Apply the greedy-split algorithm, as described in [5, Section V] and [27], to each of the joint distributions  $\alpha_{X_i, Y}^{(i)}(x', y)$ ,  $i = 1, \dots, q - 1$ , producing the joint distribution  $\beta_{X_i, Z_i, Y}^{(i)}(x', z', y)$ ;
- 3) Produce the joint distribution  $P_{X, Z, Y}^*(x, z, y)$  from the  $\beta_{X_i, Z_i, Y}^{(i)}(x', z', y)$ , according to (11) and (12), and optionally marginalize over  $y$  to produce  $P_{X, Z}^*(x, z)$ .

Our main result is the following.

*Theorem 2:* The distribution  $P_{X, Z, Y}^*(x, z, y)$  specified in (11) is a valid probability distribution, it induces a Markov chain  $X - Z - Y$  in this order, and it marginalizes to  $\sum_{z \in \mathcal{Z}} P_{X, Z, Y}^*(x, z, y) = P_{X, Y}(x, y)$ . Furthermore, under  $P_{X, Z, Y}^*(x, z, y)$  we have that

$$I(X; Z) - I(X; Y) \leq 128(q - 1) \left\lfloor L^{1/(q-1)} \right\rfloor^{-2} \quad (13)$$

The construction of  $P_{X, Z, Y}^*(x, z, y)$ , and optionally its marginalization to  $P_{X, Z}^*(x, z)$  can be done using  $O(q|\mathcal{Y}|(\log |\mathcal{Y}| + q \cdot 2^{q-1}))$  simple operations.

## V. PROOF OF THEOREM 2

### A. Informal Explanation

In this subsection we give an intuitive reasoning to why the proposed construction, i.e., the joint distribution on  $(X, Z, Y)$

specified by (12), indeed induces the correct marginal distribution on  $(X, Y)$ , satisfies the required Markov relation  $X - Z - Y$ , and attains a small mutual information gap  $I(X; Z) - I(X; Y)$ .

We begin, by writing the joint distribution on  $(X, Y)$  as

$$\begin{aligned} P_{X,Y}(x, y) &= P_Y(y) P_{X_1, \dots, X_{q-1}|Y}(x_1, \dots, x_{q-1}|y) \\ &= P_Y(y) \prod_{i=1}^{q-1} P_{X_i|Y, X_1^{i-1}}(x_i|y, x_1^{i-1}). \end{aligned}$$

Next, note that by definition of our one-hot encoding, for all  $x_1^{i-1} \neq 0_1^{i-1}$  and  $y \in \mathcal{Y}$ , we have that

$$P_{X_i|Y, X_1^{i-1}}(x_i|y, x_1^{i-1}) = \begin{cases} 0 & x_i = 1 \\ 1 & x_i = 0 \end{cases}. \quad (14)$$

It follows that we can “simulate” the channel  $P_{X_i|Y, X_1^{i-1}} : \mathcal{Y} \times \{0, 1\}^{i-1} \rightarrow \{0, 1\}$  by first passing  $Y$  through the channel  $P_{X_i|Y, X_1^{i-1}=0_1^{i-1}} : \mathcal{Y} \rightarrow \{0, 1\}$ , and then multiplying the result, which we denote  $\tilde{X}_i$ , by  $\mathbb{1}_{\{X_1^{i-1}=0_1^{i-1}\}}$ . Recalling that by (8), the channels  $P_{X_i|Y, X_1^{i-1}=0_1^{i-1}}$  are precisely the channels  $\alpha_{X_i|Y}^{(i)}$ , and noting that the events  $\{X_1^{i-1} = 0_1^{i-1}\}$  are equivalent to  $\{\tilde{X}_1^{i-1} = 0_1^{i-1}\}$ , for all  $i = 1, \dots, q-1$ , we see that

$$\begin{aligned} P_{X,Y}(x, y) &= P_Y(y) \\ &\cdot \sum_{\tilde{x}^{q-1} \in \{0,1\}^{q-1}} \prod_{i=1}^{q-1} \alpha_{X_i|Y}^{(i)}(\tilde{x}_i|y) \prod_{i=1}^{q-1} \mathbb{1}_{\{x_i=f_i(\tilde{x}_i)\}}, \end{aligned} \quad (15)$$

where  $f_i(\tilde{x}_1^i) \triangleq \tilde{x}_i \cdot \mathbb{1}_{\{\tilde{x}_1^{i-1}=0_1^{i-1}\}}$ . This distribution corresponds to generating  $Y \sim P_Y$ , then generating  $\tilde{X}_1^{q-1}$  by passing  $Y$  through the product channel  $\prod_{i=1}^{q-1} \alpha_{X_i|Y}^{(i)}$  and then generating  $X_1^{q-1}$  by the deterministic transformation

$$(X_1, \dots, X_{q-1}) = (f_1(\tilde{X}_1), \dots, f_{q-1}(\tilde{X}_1^{q-1})).$$

This view of the generation process of  $(X, Y)$  is illustrated in Figure 1.

Recall that by (10), for all  $i = 1, \dots, q-1$ , the channel  $\alpha_{X_i|Y}^{(i)}$  is equivalent to the concatenation of the channels  $\beta_{Z_i|Y}^{(i)}$  and  $\beta_{X_i|Z_i}^{(i)}$ . It therefore immediately follows that the joint distribution on  $(Y, Z_1^{q-1}, X)$  depicted in Figure 2 induces the same marginal distribution on  $(X, Y)$  as  $P_{X,Y}(x, y)$ . Furthermore, under the distribution depicted in Figure 2, the Markov relation  $Y - Z_1^{q-1} - \tilde{X}_1^{q-1} - X$  clearly holds, and consequently, so do the required Markov relation  $Y - Z_1^{q-1} - X$ . Observing that the distribution depicted in Figure 2 is precisely the one prescribed in (11), we conclude that (11) is indeed a valid distribution for the upgradation problem.

To quantify the increase of mutual information due to this upgrading procedure, we write

$$\begin{aligned} I(X; Z) &= I(X_1^{q-1}; Z_1^{q-1}) \\ &= \sum_{i=1}^{q-1} I(X_i; Z_1^{q-1} | X_1^{i-1}) \\ &= \sum_{i=1}^{q-1} \mathbb{P}(X_1^{i-1} = 0_1^{i-1}) I(X_i; Z_1^{q-1} | X_1^{i-1} = 0_1^{i-1}) \quad (16) \\ &= \sum_{i=1}^{q-1} \mathbb{P}(X_1^{i-1} = 0_1^{i-1}) I(\tilde{X}_i; Z_1^{q-1} | X_1^{i-1} = 0_1^{i-1}), \quad (17) \end{aligned}$$

where (16) follows since  $X_i$  is deterministically equal to 0 unless  $X_1^{i-1} = 0_1^{i-1}$ , and (17) follows since  $\tilde{X}_i = X_i$  whenever  $X_1^{i-1} = 0_1^{i-1}$ . Next, we observe that by construction of the distribution, the Markov chain  $\tilde{X}_i - Z_i - (Z_{\sim i}, X_1^{i-1})$  holds, where  $Z_{\sim i} = (Z_1, Z_2, \dots, Z_{i-1}, Z_{i+1}, Z_{i+2}, \dots, Z_{q-1})$ . This implies that  $\tilde{X}_i - (Z_i, X_1^{i-1}) - Z_{\sim i}$  form a Markov chain in this order, such that

$$I(\tilde{X}_i; Z_{\sim i} | Z_i, X_1^{i-1}) = 0,$$

and in particular

$$\mathbb{P}(X_1^{i-1} = 0_1^{i-1}) I(\tilde{X}_i; Z_{\sim i} | Z_i, X_1^{i-1} = 0_1^{i-1}) = 0. \quad (18)$$

Substituting (18) into (17), we get

$$I(X; Z) \leq \sum_{i=1}^{q-1} \mathbb{P}(X_1^{i-1} = 0_1^{i-1}) I(\tilde{X}_i; Z_i | X_1^{i-1} = 0_1^{i-1}).$$

Now, recalling that by construction of  $\beta_{X_i, Z_i, Y}^{(i)}(x', z', y)$

$$\begin{aligned} I(\tilde{X}_i; Z_i | X_1^{i-1} = 0_1^{i-1}) &\leq I(\tilde{X}_i; Y | X_1^{i-1} = 0_1^{i-1}) + 128\Lambda^{-2} \\ &= I(X_i; Y | X_1^{i-1} = 0_1^{i-1}) + 128\Lambda^{-2} \end{aligned}$$

and noting that

$$I(X; Y) = \sum_{i=1}^{q-1} \mathbb{P}(X_1^{i-1} = 0_1^{i-1}) I(X_i; Y | X_1^{i-1} = 0_1^{i-1}),$$

we obtain

$$\begin{aligned} I(X; Z) - I(X; Y) &\leq \sum_{i=1}^{q-1} \mathbb{P}(X_1^{i-1} = 0_1^{i-1}) 128\Lambda^{-2} \\ &\leq 128(q-1) \left( \left[ L^{1/(q-1)} \right] \right)^{-2}. \end{aligned}$$

## B. Formal Proof

*Claim 3:* The function  $P_{X,Z,Y}^*$  defined in (11) is a valid probability distribution. Specifically, summing the last parenthesized expression in (11) over all  $x \in \mathcal{X}$  yields 1; summing the first parenthesized expression in (11) over all  $z \in \mathcal{Z}$  yields 1, summing the term  $P_Y(y)$  over all  $y \in \mathcal{Y}$  yields 1.

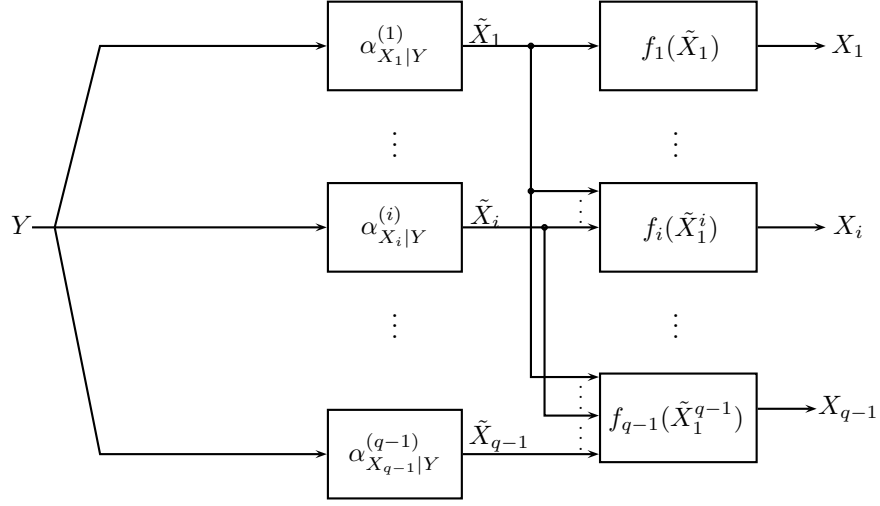


Fig. 1. A schematic illustration of  $P_{Y, X_1^{q-1}}$ , as written in (15). The functions  $f_i$  are defined as  $f_i(\tilde{x}_1^i) = \tilde{x}_1^i \cdot \mathbb{1}_{\{\tilde{x}_1^{i-1} = 0_1^{i-1}\}}$ , for  $i = 1, \dots, q-1$ .

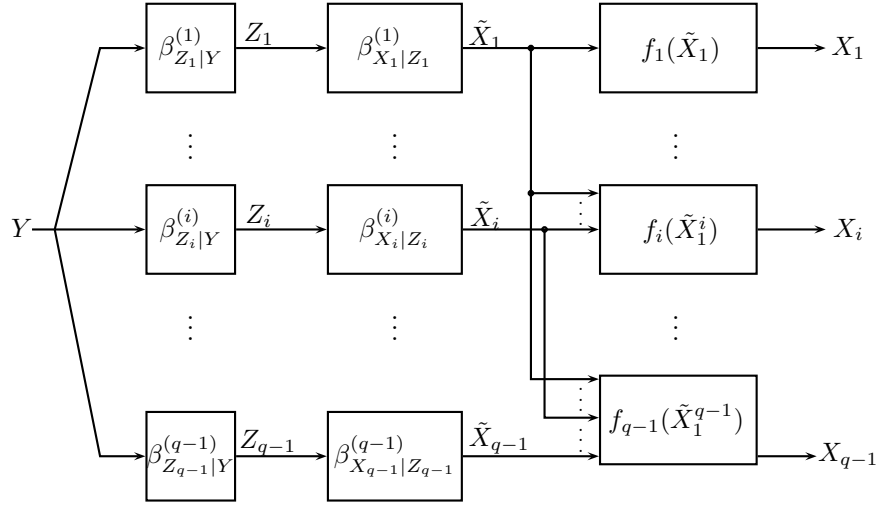


Fig. 2. A schematic illustration of the distributed construction in (11). The functions  $f_i$  are defined as  $f_i(\tilde{x}_1^i) = \tilde{x}_1^i \cdot \mathbb{1}_{\{\tilde{x}_1^{i-1} = 0_1^{i-1}\}}$ , for  $i = 1, \dots, q-1$ .

*Proof:* By inspection, all the expressions involved are non-negative. Fix some  $z \in \mathcal{Z}$  and consider the last parenthesized expression in (11). We abuse notation and write

$$\begin{aligned} & \sum_{x \in \mathcal{X}} \prod_{i=1}^{q-1} \gamma_{X_i|Z_i, X_1^{i-1}}^{(i)}(x_i|z_i, x_1^{i-1}) \\ &= \sum_{(x_1, \dots, x_{q-1}) \in (\mathcal{X}')^{q-1}} \prod_{i=1}^{q-1} \gamma_{X_i|Z_i, X_1^{i-1}}^{(i)}(x_i|z_i, x_1^{i-1}) \\ &= \prod_{i=1}^{q-1} \sum_{x_i \in \mathcal{X}'} \gamma_{X_i|Z_i, X_1^{i-1}}^{(i)}(x_i|z_i, x_1^{i-1}), \end{aligned}$$

where for the first equality we recall by (12) that vectors  $(x_1, x_2, \dots, x_{q-1})$  with support greater than 1 contribute nothing

to the above sum, and the RHS is short for

$$\begin{aligned} & \sum_{x_1 \in \mathcal{X}'} \gamma_{X_1|Z_1}^{(1)}(x_1|z_1) \sum_{x_2 \in \mathcal{X}'} \gamma_{X_2|Z_2, X_1}^{(2)}(x_2|z_2, x_1) \\ & \quad \cdots \sum_{x_{q-1} \in \mathcal{X}'} \gamma_{X_{q-1}|Z_{q-1}, X_1^{q-2}}^{(q-1)}(x_{q-1}|z_{q-1}, x_1^{q-2}). \end{aligned}$$

Starting from the innermost sum ( $i = q-1$ ) and working out, and recalling the definition of  $\gamma^{(i)}$  in (12), we see that the above equals 1.

The sub-claim about the first parenthesized quantity in (11) is proved similarly. The sub-claim about summing  $P_Y(y)$  over all  $y \in \mathcal{Y}$  follows trivially, by virtue of  $P_Y$  being a probability distribution.  $\blacksquare$

*Claim 4:* Let the random variables  $X$ ,  $Z$ , and  $Y$  be defined by the distribution  $P_{X, Z, Y}^*$  given in (11). Then,  $X - Z - Y$  form a Markov chain in this order.

*Proof:* We must show that  $P_{X, Z, Y}^*$  can be factored as in (6), where each term in this factor is a valid probability

distribution. By Claim 3 and inspection of (11), this is indeed the case. ■

*Claim 5:* Marginalizing the distribution  $P_{X,Z,Y}^*$  defined in (11) over all  $z \in \mathcal{Z}$  results in the original joint distribution  $P_{X,Y}$ .

*Proof:* Fix some  $x \in \mathcal{X}$  and  $y \in \mathcal{Y}$ . We must show that

$$P_{X,Y}(x,y) = \sum_{z \in \mathcal{Z}} P_{X,Z,Y}^*(x,z,y).$$

By inspection of (11), this is equivalent to proving that

$$\begin{aligned} P_{X|Y}(x|y) &= \sum_{z \in \mathcal{Z}} \prod_{i=1}^{q-1} \beta_{Z_i|Y}^{(i)}(z_i|y) \gamma_{X_i|Z_i, X_1^{i-1}}^{(i)}(x_i|z_i, x_1^{i-1}). \end{aligned} \quad (19)$$

As before, we may simplify the RHS to

$$\prod_{i=1}^{q-1} \sum_{z_i \in \mathcal{Z}_i} \beta_{Z_i|Y}^{(i)}(z_i|y) \gamma_{X_i|Z_i, X_1^{i-1}}^{(i)}(x_i|z_i, x_1^{i-1}). \quad (20)$$

Recall that  $1 \leq x \leq q$ . Consider first a term  $i$  in the product, where  $i > x$ . In this case  $x_1^{i-1}$  is non-zero, and hence  $\gamma_{X_i|Z_i, X_1^{i-1}}^{(i)}(x_i|z_i, x_1^{i-1})$  simply equals 1, by (12). Hence, term  $i$  in (20) is simply

$$\sum_{z_i \in \mathcal{Z}_i} \beta_{Z_i|Y}^{(i)}(z_i|y) = 1.$$

Now consider a term  $i$  in the product (20) for which  $i \leq x$ . In this case,  $x_1^{i-1}$  is the zero-vector of length  $i-1$ , and hence  $\gamma_{X_i|Z_i, X_1^{i-1}}^{(i)}(x_i|z_i, x_1^{i-1})$  equals  $\beta_{X_i|Z_i}^{(i)}(x_i|z_i)$ , by (12). Hence, term  $i$  in (20) is

$$\sum_{z_i \in \mathcal{Z}_i} \beta_{Z_i|Y}^{(i)}(z_i|y) \beta_{X_i|Z_i}^{(i)}(x_i|z_i) = \alpha^{(i)}(x_i|y),$$

by virtue of the upgrading process having the marginalization property. Combining these two observations with the definition of  $\alpha^{(i)}$  in (8), we simplify (20) to

$$\prod_{i=1}^{\min\{x, q-1\}} P(X_i = x_i | Y = y, X_1^{i-1} = x_1^{i-1}),$$

where the probabilities are calculated according to the given probability distribution  $P_{X,Y}$ . Recalling the one-hot convention, we see that for both the case  $x = q$  as well as the case  $x < q$ , the above indeed equals the LHS of (19). ■

*Claim 6:* Let  $X$ ,  $Z$ , and  $Y$  be the random variables corresponding to the distribution  $P_{X,Z,Y}^*$  defined in (11). Fix  $0 \leq j \leq q-1$ . Then, for  $x_j \in \mathcal{X}'$ ,  $z_j \in \mathcal{Z}_j$ , and  $y \in \mathcal{Y}$ , We have

$$\begin{aligned} \mathbb{P}(X_j = x_j, Z_j = z_j, Y = y | X_1^{j-1} = 0_1^{j-1}) \\ = \beta^{(j)}(x_j, z_j, y). \end{aligned} \quad (21)$$

*Proof:* It suffices to prove that

$$\begin{aligned} \mathbb{P}(X_j = x_j, X_1^{j-1} = 0_1^{j-1}, Z_j = z_j, Y = y) \\ = \beta^{(j)}(x_j, z_j, y) \mathbb{P}(X_1^{j-1} = 0_1^{j-1}). \end{aligned} \quad (22)$$

We get the LHS of (22) by fixing  $x_1^{j-1} = 0_1^{j-1}$ , and summing  $P^*(x, z, y)$  over all  $x_{j+1}^{q-1}$ ,  $z_{j+1}^q$ , and  $z_1^{j-1}$ . Summing (11) over only the first two terms,  $x_{j+1}^{q-1}$  and  $z_{j+1}^q$ , causes both products in (11) to be from 1 to  $j$ . Also, since  $x_1^{j-1} = 0^{j-1}$ , we get from (12) that the  $\gamma^{(i)}$  term in the second product of (11) can be replaced by  $\beta_{X_i|Z_i}^{(i)}(x_i|z_i)$ . Thus, we have,

$$\begin{aligned} \sum_{x_{j+1}^{q-1}, z_{j+1}^q} P_{X,Z,Y}^*(x, z, y) \\ = P_Y(y) \cdot \left( \prod_{i=1}^j \beta_{Z_i|Y}^{(i)}(z_i|y) \cdot \beta_{X_i|Z_i}^{(i)}(x_i|z_i) \right) \\ = P_Y(y) \cdot \left( \prod_{i=1}^j \beta_{X_i, Z_i|Y}^{(i)}(x_i, z_i|y) \right), \end{aligned} \quad (23)$$

where the second equality follows from the Markovity promised by our upgrading procedure. Recall that we have yet to sum over all  $z_1^{j-1}$ . Doing so causes all terms on the RHS of (23), save for the term  $i = j$  to be marginalized to  $\beta_{X_i|Y}^{(i)}(x_i|y)$ . Now we recall that by the definition of upgrading,  $\beta_{X_i|Y}^{(i)}(x_i|y) = \alpha_{X_i|Y}^{(i)}(x_i|y)$ , and by the definition of  $\alpha^{(i)}$  in (8) we conclude that

$$\begin{aligned} \mathbb{P}(X_j = x_j, X_1^{j-1} = 0_1^{j-1}, Z_j = z_j, Y = y) \\ = \sum_{x_{j+1}^{q-1}, z_{j+1}^q, z_1^{j-1}} P_{X,Z,Y}^*(x, z, y) \\ = P_Y(y) \cdot \left( \prod_{i=1}^{j-1} \mathbb{P}(X_i = 0 | Y = y, X_1^{i-1} = 0_1^{i-1}) \right) \\ \cdot \beta_{X_j, Z_j|Y}^{(j)}(x_j, z_j|y), \end{aligned} \quad (24)$$

where the probabilities  $\mathbb{P}(\cdot)$  above are calculated according to the given probability distribution  $P_{X,Y}$ , by virtue of this being the probability distribution through which the  $\alpha^{(i)}$  are defined. We now note that the RHS of (24) can be simplified to

$$\mathbb{P}(Y = y, X_1^{j-1} = 0_1^{j-1}) \cdot \beta_{X_j, Z_j|Y}^{(j)}(x_j, z_j|y),$$

where again, the  $\mathbb{P}$  are according to the given probability distribution  $P_{X,Y}$ . Thus,

$$\begin{aligned} \mathbb{P}(X_j = x_j, X_1^{j-1} = 0_1^{j-1}, Z_j = z_j, Y = y) \\ = \mathbb{P}(Y = y, X_1^{j-1} = 0_1^{j-1}) \cdot \beta_{X_j, Z_j|Y}^{(j)}(x_j, z_j|y) \\ = \mathbb{P}(X_1^{j-1} = 0_1^{j-1}) \\ \cdot \mathbb{P}(Y = y | X_1^{j-1} = 0_1^{j-1}) \cdot \beta_{X_j, Z_j|Y}^{(j)}(x_j, z_j|y) \\ = \mathbb{P}(X_1^{j-1} = 0_1^{j-1}) \cdot \alpha_Y^{(j)}(y) \cdot \beta_{X_j, Z_j|Y}^{(j)}(x_j, z_j|y) \\ = \mathbb{P}(X_1^{j-1} = 0_1^{j-1}) \cdot \beta_Y^{(j)}(y) \cdot \beta_{X_j, Z_j|Y}^{(j)}(x_j, z_j|y) \\ = \mathbb{P}(X_1^{j-1} = 0_1^{j-1}) \cdot \beta_{X_j, Z_j, Y}^{(j)}(x_j, z_j, y), \end{aligned}$$

proving (22). ■

Now that we have established the  $P_{X,Z,Y}^*$  is a valid probability distribution, in the sense of upgrading  $P_{X,Y}$ , we give a bound on the upgrading performance. That is, define the random variables  $X$ ,  $Z$ , and  $Y$  according to  $P_{X,Z,Y}^*$ . We now bound  $H(X|Y) - H(X|Z)$  from above.

*Lemma 7:* Let a joint distribution  $P_{X,Y}$  and a parameter  $L$  be given. Construct  $P_{X,Z,Y}^*$  be as defined in (11). Let the random variables  $X$ ,  $Z$ , and  $Y$  be defined according to  $P_{X,Z,Y}^*$ . Then,

$$H(X|Y) - H(X|Z) \leq 128 \cdot (|\mathcal{X}| - 1) \cdot \Lambda^{-2}. \quad (25)$$

*Proof:* Recall that our figure of merit is  $I(X; Z|Y)$ , by (5). From the chain rule,

$$\begin{aligned} I(X; Z|Y) &= I(X_1, X_2, \dots, X_{q-1}; Z|Y) \\ &= \sum_{i=1}^{q-1} I(X_i; Z|Y, X_1^{i-1}) \\ &\stackrel{(a)}{=} \sum_{i=1}^{q-1} P(X_1^{i-1} = 0_1^{i-1}) \cdot I(X_i; Z|Y, X_1^{i-1} = 0_1^{i-1}) \\ &\leq \sum_{i=1}^{q-1} I(X_i; Z|Y, X_1^{i-1} = 0_1^{i-1}), \end{aligned}$$

where (a) follows from the one-hot representation: if  $X_1^{i-1} \neq 0_1^{i-1}$ , then  $X_i$  is a degenerate random variable always equal to 0. It would be easy to bound the term  $I(X_i; Z|Y, X_1^{i-1} = 0_1^{i-1})$ , if  $Z$  were replaced by  $Z_i$ . Namely, Claim 6 and our binary-input upgrading algorithm ensures that<sup>6</sup>

$$I(X_i; Z_i|Y, X_1^{i-1} = 0_1^{i-1}) \leq 128 \cdot \Lambda^{-2}.$$

Thus, our result will be proved once we show that

$$I(X_i; Z|Y, X_1^{i-1} = 0_1^{i-1}) = I(X_i; Z_i|Y, X_1^{i-1} = 0_1^{i-1}).$$

By the chain rule, this is equivalent to showing that

$$I(X_i; Z_{\sim i}|Y, Z_i, X_1^{i-1} = 0_1^{i-1}) = 0,$$

where  $Z_{\sim i} = (Z_1, Z_2, \dots, Z_{i-1}, Z_{i+1}, Z_{i+2}, \dots, Z_{q-1})$ , as defined above. That is, we must show that  $X_i$  and  $Z_{\sim i}$  are independent, when conditioning on an event  $A$  of the form

$$A = \{Y = y, Z_i = z_i, X_1^{i-1} = 0_1^{i-1}\}.$$

Hence, fix  $y \in \mathcal{Y}$ ,  $x_i \in \mathcal{X}$ , and  $z \in \mathcal{Z}$ , and let us show that

$$\begin{aligned} \mathbb{P}(A) \cdot \mathbb{P}(A, X_i = x_i, Z_{\sim i} = z_{\sim i}) \\ = \mathbb{P}(A, X_i = x_i) \cdot \mathbb{P}(A, Z_{\sim i} = z_{\sim i}). \end{aligned} \quad (26)$$

<sup>6</sup>This is only guaranteed for  $\Lambda \geq 2$ , but since  $128 \cdot \Lambda^{-2} > \log 2$  otherwise, we may assume  $\Lambda \geq 2$  without loss of generality.

We now use (11) and (12) to write each term of (26) explicitly. Namely, one easily gets that

$$\begin{aligned} \mathbb{P}(A) &= P_Y(y) \cdot \beta_{Z_i|Y}^{(i)}(z_i|y) \prod_{j=1}^{i-1} \beta_{X_j|Y}^{(j)}(0|y) \\ \mathbb{P}(A, X_i = x_i, Z_{\sim i} = z_{\sim i}) &= P_Y(y) \cdot \beta_{X_i, Z_i|Y}^{(i)}(x_i, z_i|y) \prod_{j=1}^{i-1} \beta_{X_j|Y}^{(j)}(0|y) \prod_{j=i+1}^{q-1} \beta_{Z_j|Y}^{(j)}(z_j|y) \\ \mathbb{P}(A, X_i = x_i) &= P_Y(y) \cdot \beta_{X_i, Z_i|Y}^{(i)}(x_i, z_i|y) \prod_{j=1}^{i-1} \beta_{X_j|Y}^{(j)}(0|y) \\ \mathbb{P}(A, Z_{\sim i} = z_{\sim i}) &= P_Y(y) \cdot \beta_{Z_i|Y}^{(i)}(z_i|y) \prod_{j=1}^{i-1} \beta_{X_j|Y}^{(j)}(0|y) \prod_{j=i+1}^{q-1} \beta_{Z_j|Y}^{(j)}(z_j|y). \end{aligned}$$

Using the above, we easily verify (26).  $\blacksquare$

*Proof of Theorem 2:* By Claims 4 and 5, we have indeed constructed an upgrading of the original distribution. Lemma 7 ensures that the difference in entropies satisfies (13), by (9) and (25). Also, the output alphabet size is at most  $\Lambda^{q-1} \leq L$ , which follows by recalling that  $Z = Z_1^{q-1}$  and the definition of  $\Lambda$  in (9).

All that remains is to discuss the complexity of our algorithm. The construction of the distributions  $\alpha^{(i)}$  given in (8) for  $1 \leq i \leq q-1$  is easy to derive, if for each  $y \in \mathcal{Y}$  and  $0 \leq i \leq q-1$  we have the probabilities  $\mathbb{P}(Y = y, X_1^i = 0^i)$  at hand. Next, we note that  $\mathbb{P}(Y = y, X_1^i = 0^i)$  is readily computed from  $\mathbb{P}(Y = y, X_1^{i+1} = 0^{i+1})$ . Thus, the calculation of all of the  $\alpha^{(i)}$  takes time  $O(q \cdot |\mathcal{Y}|)$ .

In order to calculate the distributions  $\beta^{(i)}$  for  $1 \leq i \leq q-1$ , the binary upgrading algorithm is run, taking time  $O(|\mathcal{Y}| \log |\mathcal{Y}|)$  for each of the  $q-1$  invocations. Thus, the time needed to construct the joint distribution  $P_{X,Z,Y}^*$  is  $O(q \cdot |\mathcal{Y}| \log |\mathcal{Y}|)$ .

Let us now discuss the complexity of marginalizing  $P_{X,Z,Y}^*$  to produce  $P_{X,Z}^*$ . Clearly, this can be accomplished in time  $O(q \cdot |\mathcal{Y}| \cdot |\mathcal{Z}|)$ . However, we can do better by first noting that even though we produce, for  $1 \leq i \leq q-1$ , a joint probability  $\beta^{(i)}(x', y, z_i)$  involving three variables:  $x' \in \mathcal{X}'$  binary,  $y \in \mathcal{Y}$ , and  $z \in \mathcal{Z}^{(i)}$ , this probability distribution is very sparse. Namely, for each  $y \in \mathcal{Y}$  there are at most two  $z_i \in \mathcal{Z}^{(i)}$  such that  $\beta^{(i)}(y, z_i) > 0$ . This is proved by induction on the number of upgrading steps in the binary upgrading algorithm. The key observation is that a symbol which is removed from the output alphabet due to upgrading is ‘split’ between two neighboring symbols, as can be seen in (42) below.

As the first step of our efficient marginalization, for each  $1 \leq i \leq q-1$  and  $y \in \mathcal{Y}$ , let us build the subset  $\mathcal{Z}^{(i)}(y)$  of  $\mathcal{Z}^{(i)}$  for which the probability  $\beta^{(i)}(y, z_i)$  is positive if and only if  $z_i \in \mathcal{Z}^{(i)}(y)$ . As explained earlier, the size of such a set  $\mathcal{Z}^{(i)}(y)$  is always at most 2. Also, the time needed to construct these sets is  $O(q \cdot |\mathcal{Y}|)$ , if the binary upgrading algorithm is modified to retain the relevant information.



Next, in order to calculate  $P_{X,Z}^*$ , we define an array indexed by  $x \in \mathcal{X}$  and  $z = (z^{(1)}, z^{(2)}, \dots, z^{(q-1)}) \in \mathcal{Z}$ . All entries of the array are initialized to 0. Then, we have an outermost loop on all  $y \in \mathcal{Y}$ , a mid-level loop on all  $z = (z^{(1)}, z^{(2)}, \dots, z^{(q-1)})$  such that  $z^{(i)} \in \mathcal{Z}^{(i)}(y)$  for all  $1 \leq i \leq q-1$ , and an inner-most loop on  $x$ , going from 1 up to  $q$ . The operation carried out in the innermost loop is adding  $P_{X,Z,Y}^*(x, z, y)$  to entry  $(x, y)$  of our table. Clearly, when the calculation finishes, entry  $(x, z)$  of our table equals  $P_{X,Z}^*(x, z)$ . Note also that by (11) and (12), we can share calculations between different values of  $x$ , such that time needed for the inner most loop to cycle over all  $x \in \mathcal{X}$  is  $O(q)$ . It follows that the cost of marginalizing  $P_{X,Z,Y}^*$  to produce  $P_{X,Z}^*$  can be accomplished in time  $O(q \cdot 2^{q-1} \cdot |\mathcal{Y}|)$ . ■

## VI. AN IMPROVED CHANNEL DEGRADING ALGORITHM

While the focus of this paper is on channel upgrading, we now argue that similar ideas can also be used for channel degrading. Recall that a channel  $W$  is degraded with respect to a channel  $Q$  if  $Q$  is upgraded with respect to  $W$ .

Although the optimal power law for channel degradation was already established in [27], we now provide a channel degrading algorithm that attains the same power law with better (smaller) constants, and more importantly, with reduced complexity compared to that of the algorithm analyzed in [27].

The algorithm studied in [27] is the so-called ‘greedy-merge’ algorithm, which was introduced before under various different names [5], [25], [32]. The algorithm starts with a joint distribution  $P_{X,Y}$  supported on  $\mathcal{X} \times \mathcal{Y}$  and outputs a joint distribution  $P_{X,Z}$  supported on  $\mathcal{X} \times \mathcal{Z}$ , where  $|\mathcal{Z}| = L < |\mathcal{Y}|$ , and the channel  $P_{Z|X}$  is degraded with respect to  $P_{Y|X}$ . The algorithm can also output the function  $f$  that maps (quantizes) symbols from  $\mathcal{Y}$  to  $\mathcal{Z}$ . The algorithm implements an iterative process that merges two symbols of the output alphabet at each step, hence, reducing the size of the output alphabet by one letter. The two symbols that are merged at each step are those whose merging reduces the mutual information between  $X$  and the channel output by the smallest amount. In general, one must enumerate over all possible pairs before deciding which pair to merge, leading to a computational complexity of at least  $\Omega(|\mathcal{Y}|^2)$  operations, see [25]. However, in the special case where  $|\mathcal{X}| = 2$ , it can be shown (see [5, Theorem 8] and Remark 16 in the appendix) that the optimal pair can be found by relabeling the  $|\mathcal{Y}|$  output symbols such that  $p_1 \leq \dots \leq p_{|\mathcal{Y}|}$ , where  $p_i = \mathbb{P}(X = 0|Y = i)$ , and then finding the best pair for merging among the pairs  $(i, i+1)$ ,  $i = 1, \dots, |\mathcal{Y}| - 1$ . Thus, in the case where  $\mathcal{X}$  is binary, the computational complexity of a single merge is reduced to  $O(|\mathcal{Y}| \log |\mathcal{Y}|)$ . Further noting that after every merge, only two new pairs of adjacent letters arise, we can save many computations, and implement all  $|\mathcal{Y}| - L$  iterations at a total computational cost of  $O(|\mathcal{Y}| \log |\mathcal{Y}|)$  operations.

Based on the one-hot representation of  $X$ , it was shown in [28] that one can leverage any degrading algorithm with optimal power law for  $|\mathcal{X}| = 2$  to an optimal-power-law degrading algorithm for general alphabet size. However, an explicit algorithm and an accompanying computational com-

plexity analysis was not given there. Below, we explicitly specify such an algorithm and then analyze its performance.

Let  $\Lambda \leq |\mathcal{Y}|$  be some positive integer. Our algorithm degrades the channel  $P_{Y|X} : \{1, \dots, q\} \rightarrow \mathcal{Y}$  to the channel  $P_{f(Y)|X} : \{1, \dots, q\} \rightarrow \{1, \dots, L\}$ , where  $f : \mathcal{Y} \rightarrow \{1, \dots, L\}$  is a deterministic function (quantizer) with cardinality  $L \leq \Lambda^{q-1}$ . The algorithm consists of three steps:

- 1) Construct the distributions  $\alpha_{X_i, Y}^{(i)}(x', y)$  for  $i = 1, \dots, q-1$ , according to (8). Each such distribution is supported on  $\{0, 1\} \times \mathcal{Y}$ ;
- 2) Apply the greedy-merge algorithm, as described in [5, Algorithm C] and [27], for each of the joint distributions  $\alpha_{X_i, Y}^{(i)}(x', y)$ ,  $i = 1, \dots, q-1$ , to degrade the channel  $\alpha_{Y|X_i}^{(i)}(y|x') = \mathbb{P}(Y = y|X_i = x', X_1^{i-1} = 0_1^{i-1})$  to the channel  $P_{f_i(Y)|X_i}$  with  $f_i : \mathcal{Y} \rightarrow \{1, \dots, \Lambda\}$ .
- 3) Set  $f : \mathcal{Y} \rightarrow \{1, \dots, \Lambda\}^{q-1}$  as  $f(y) = [f_1(y), \dots, f_{q-1}(y)]$  and compute the resulting joint distribution  $P_{X, f(Y)}$  as

$$P_{X, f(Y)}(x, t_1^{q-1}) = \sum_{y \in \mathcal{Y}} P_{X, Y}(x, y) \prod_{i=1}^{q-1} \mathbb{1}_{\{f_i(y) = t_i\}},$$

for all  $(x, t_1^{q-1}) \in \{1, \dots, q\} \times \{1, \dots, \Lambda\}^{q-1}$ .

*Theorem 8:* Let  $(X, Y) \sim P_{X, Y}$  be random variables in  $\mathcal{X} \times \mathcal{Y}$ , where  $|\mathcal{X}| = q > 1$ , and  $\mathcal{Y}$  is discrete. For any integer  $L$ , the algorithm defined above with  $\Lambda = \lfloor L^{1/(q-1)} \rfloor$  satisfies

$$I(X; Y) - I(X; f(Y)) \leq 64(q-1) \cdot \left[ L^{1/(q-1)} \right]^{-2}. \quad (27)$$

The construction of the joint distribution  $P_{X, f(Y)}$  can be done using  $O(q|\mathcal{Y}| \log |\mathcal{Y}|)$  simple operations.

*Proof:* The proof of (27) is similar to [28, Proof of Proposition 7], albeit with improved constants. The improved constants result from a tighter analysis of the degradation loss of the greedy-merge algorithm for the binary case  $|\mathcal{X}| = 2$ , which is derived in Appendix C. In Appendix D we prove (27) based on this result.

As for the time complexity: Step 1 was shown to be performed using  $O(q|\mathcal{Y}|)$  simple operations in the proof of Theorem 2. Each application of the greedy-merge algorithm for a binary input alphabet can be performed in  $O(|\mathcal{Y}| \log |\mathcal{Y}|)$  operations [5] and therefore the computational cost of step 2 is  $O(q|\mathcal{Y}| \log |\mathcal{Y}|)$ . Finally, for the computation of Step 3, we can construct a table of size  $q \times \Lambda^{q-1}$  and initialize it with zeros. Then, for each  $y \in \mathcal{Y}$ , we first calculate  $f(y) = [f_1(y), \dots, f_{q-1}(y)]$ , which can be done in time  $O(q)$  for each  $y$ , since each of the  $f_i$  can be implemented using an array (lookup table). Thus, we have calculated which column to update in our table. We now loop over all  $x$  and increase the value of the cell indexed by  $(x, f(y))$  by  $P_{X, Y}(x, y)$ . Once we have gone through all  $y \in \mathcal{Y}$  the table holds  $P_{X, f(Y)}$ . Thus, Step 3 can be performed using  $O(q|\mathcal{Y}|)$  simple operations. ■

## VII. NUMERICAL RESULTS

We now discuss some numerical results that demonstrate the effectiveness of our method. First, consider Figure 3. In it, we fix the input alphabet size to  $q = 3$  and give four plots

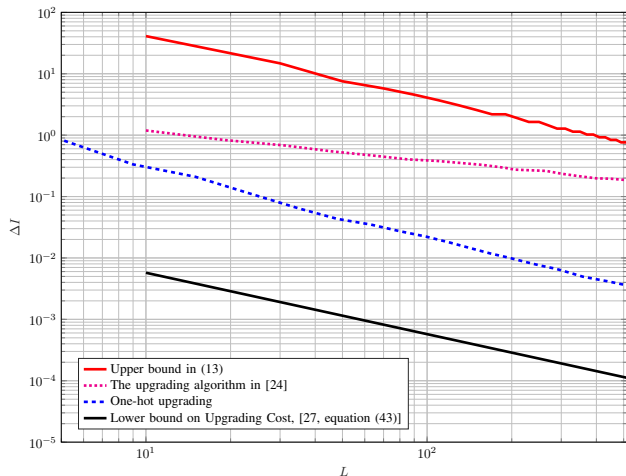


Fig. 3. **Upgrading.** Log-log plot of the difference in mutual information as a function of output alphabet size. For input alphabet size  $q = 3$ , we plot, from top to bottom: the upper bound in (13); the algorithm from [24]; our one-hot upgrading algorithm; the lower bound on Upgrading Cost given in [27, equation (43)].

of  $\Delta I$ , the difference in mutual information resulting from upgrading, as a function of  $L$ , the output alphabet size of the upgraded channel. The topmost plot is the upper bound (13) we have derived for our one-hot upgrading algorithm. The second and third plots from the top are the result of running the upgrading algorithm in [24] and our one-hot upgrading algorithm, respectively. The two algorithms were run on the channel described<sup>7</sup> in [26, Subsection V.A], with parameter  $M = 400$  and uniform input distribution. As discussed in [27, Section IV], this channel is “hard to upgrade”. We stress that for these two plots, the  $x$ -axis is in fact  $L'$ , the number of output letters in the resulting upgraded channel, *not* the “designed” output alphabet size  $L$  which is the input to these algorithms. The bottom-most plot is the lower bound on the Upgrading Cost, given in [27, equation (43)]. First, note that all plots save the second from the top have roughly the same slope. This is because the two bounds have the same power law,  $L^{-2/(q-1)}$ , and we have chosen to upgrade a “hard” channel for which this power law is mandated, and indeed achieved by our algorithm. In contrast, we can see that the slope corresponding to the upgrading algorithm in [24] is roughly half of that as ours. This is explained by noting that [24] only promises a power law of  $L^{-1/(q-1)}$ .

Figure 4 is the analog of Figure 3, but now we consider degrading. Again, we have four plots. From top to bottom, these are: the upper bound in (27) on the one-hot degrading algorithm; the result of running the algorithm in [23] on the same channel as before (which [26] proves is “hard to degrade”); running the one-hot degrading algorithm (originally presented in [28], and discussed in our Section VI) on this

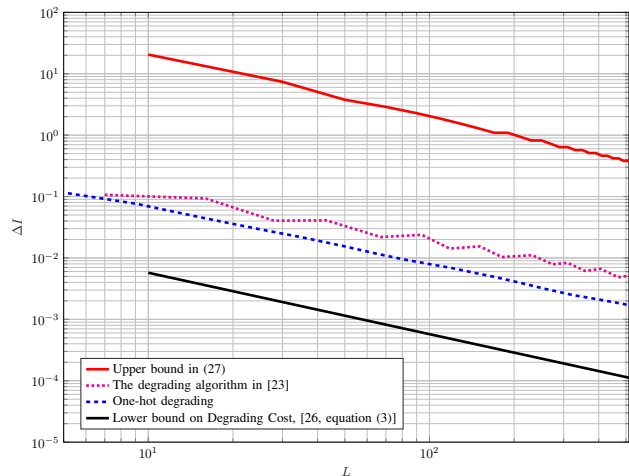


Fig. 4. **Degrading.** Log-log plot of the difference in mutual information as a function of output alphabet size. For input alphabet size  $q = 3$ , we plot, from top to bottom: the upper bound in (27); the algorithm from [23]; the one-hot degrading algorithm; the lower bound on Degrading Cost given in [26, equation (3)].

channel; the lower bound on the Degrading Cost given in [26, equation (3)]. Note that the one-hot degrading algorithm is indeed better than the prior art in [23], but this is not as pronounced as it was in the upgrading case. We conjecture that this is because the power law of the algorithm in [23] can be strengthened to something better than  $L^{-1/(q-1)}$  (proven in [24, Lemma 8]).

We close this section with another comparison between our upgrading algorithm and that of [24]. As explained in the introduction, the upgrading transform is used in order to construct polar codes whose input distribution is non-symmetric. In light of this, consider Table I. In it, each column corresponds to a different input distribution on the symbols of  $\mathcal{X} = \{0, 1, 2\}$ . That is the second column corresponds to an input distribution in which the symbol 0 has an a priori probability of 0.8, and the two symbols 1 and 2 each have an a priori probability of 0.1. The second row in the table is the entropy corresponding to each input distribution. The rest of the table is divided into two parts: the first part corresponds to the one-hot upgrading algorithm, and the second part to the upgrading algorithm from [24]. For both algorithms, the designed output alphabet size is 100. That is, for the one-hot algorithm we take  $L = 100$  and for the algorithm in [24] we take  $\mu$  such that  $[0, 1]$  is comprised of 10 regions. In both parts, we consider building a code of length 256, obtained by applying 8 polar transforms. The first row in each block is the average entropy of  $H(U_i|U^{i-1})$ , where  $1 \leq i \leq 256$ . Ideally, this average entropy should equal that of underlying input distribution. However, do to the quantization effect of upgrading, the average entropy is lower. We see, again, that the one-hot algorithm results in a much smaller reduction of entropy. The next two rows count the number of indices  $i$  for which the total variation parameter  $K(U_i|U^{i-1})$  is smaller than a prescribed number, which is 0.01 for the penultimate row in each part and 0.001 for the last row in each part. The quantity  $K(U_i|U^{i-1})$  is as defined in [15, Appendix B].

<sup>7</sup>In brief: the channel has an input alphabet of size  $q$  and an output alphabet of size  $\binom{M+q-1}{q-1}$ , where  $M$  is a parameter. The output alphabet consists of all symbols of the form  $\langle j_1, j_2, \dots, j_q \rangle$ , where the  $j_x$  are non-negative integers that sum to  $M$ . The probability of receiving  $\langle j_1, j_2, \dots, j_q \rangle$  given that  $x$  was transmitted equals  $\frac{q \cdot j_x}{M \binom{M+q-1}{q-1}}$ .

TABLE I  
COMPARISON OF ONE-HOT UPGRADING AND [24].

$(p_0, p_1, p_2)$	(0.8, 0.1, 0.1)	(0.6, 0.2, 0.2)	(0.4, 0.3, 0.3)	(0.34, 0.33, 0.33)
$H$	0.92192	1.370951	1.57095	1.58482
one-hot				
avg. $H$	0.89484	1.34842	1.56692	1.58473
@0.01	68	150	231	253
@0.001	50	129	216	247
[24]				
avg. $H$	0.50171	0.75992	1.07587	1.33681
@0.01	8	21	63	116
@0.001	8	16	33	65

We want this number to be as high as possible. That is, we count the pool of  $U_i$  with corresponding total-variation low enough to allow them to be “non-frozen” (and from this pool, we will only choose those whose Bhattacharyya parameter  $Z(U_i|U^{i-1}, Y^N)$  is low as well, in order to facilitate decoding). Again, one-hot decoding performs substantially better than [24].

A reader wishing to reproduce the above results (or obtain new ones) can use the code we have published on [33].

#### APPENDIX: UPGRADATION AND DEGRADATION IN THE BINARY CASE

The purpose of this appendix is to sharpen some of the results of [27], as well as provide simpler and shorter proofs. The main contribution of the appendix is a simple derivation of a sphere-packing bound for the simple case of  $|\mathcal{X}| = 2$ , provided in Appendix A. Then, in Appendix B and Appendix C, we essentially repeat the arguments from [27] in order to leverage the improved sphere-packing bound of Appendix A to improved upper bounds on the loss of upgradation and degradation, respectively, for the case of  $|\mathcal{X}| = 2$ . While these bounds are sharper than those reported in [27], the proofs in these subsections do not contain new ideas and are brought merely for completeness. As this paper shows, upper bounds on the loss of upgradation for this special case, immediately yield tight bounds (in terms of the power-law) for general  $|\mathcal{X}|$ . Similarly, in [28, Proposition 7] it was shown that upper bounds on the loss of degradation in the binary case yields tight power-law bounds for general  $|\mathcal{X}|$ . In Appendix D we repeat that derivation in order to prove Theorem 8 from Section VI and obtain sharper bounds on the loss for degradation for general  $|\mathcal{X}|$ .

#### APPENDIX A SPHERE-PACKING IN THE SIMPLEX OF BERNOULLI DISTRIBUTIONS

Let  $h_2(p) = -p \log(p) - (1-p) \log(1-p)$  be the binary entropy function. By the concavity of  $p \mapsto h_2(p)$  we have that  $h_2(\alpha p_0 + (1-\alpha)p_1) - \alpha h_2(p_0) - (1-\alpha)h_2(p_1) \geq 0$  for all  $\alpha, p_0, p_1 \in [0, 1]$ . The next lemma upper bounds this difference universally for all  $\alpha \in [0, 1]$ , using relatively simple functions of  $p_0, p_1$ . We remind the reader that logarithms are taken to the natural base.

*Lemma 9:* For any  $0 \leq p_0 \leq p_1 \leq 1$  and  $\alpha \in [0, 1]$ , we have that

$$\begin{aligned} & h_2(\alpha p_0 + (1-\alpha)p_1) - \alpha h_2(p_0) - (1-\alpha)h_2(p_1) \\ & \leq \min \left\{ p_1 - p_0, \frac{(p_1 - p_0)^2}{2 \min\{p_0, 1 - p_1\}} \right\}. \end{aligned}$$

We note that, up to constants, the upper bound  $p_1 - p_0$  and the upper bound  $\frac{(p_1 - p_0)^2}{2 \min\{p_0, 1 - p_1\}}$ , respectively, can be obtained by specializing [25, Lemma 1] and [27, Equation (11-12)] to the binary case.

*Proof:* Let  $A \sim \text{Ber}(1-\alpha)$  and let  $B$  be a binary random variable with conditional distributions  $[B|A=0] \sim \text{Ber}(p_0)$ , and  $[B|A=1] \sim \text{Ber}(p_1)$ . With these definitions, we have that

$$h_2(\alpha p_0 + (1-\alpha)p_1) - \alpha h_2(p_0) - (1-\alpha)h_2(p_1) = I(A; B).$$

Using the variational formula for mutual information [34, Chapter 2, Equation (3.7)], we write

$$\begin{aligned} I(A; B) &= \min_Q D(P_{B|A} \| Q | P_A) \\ &= \min_p \alpha \cdot d_2(p_0 \| p) + (1-\alpha) \cdot d_2(p_1 \| p), \end{aligned} \quad (28)$$

where  $D(P \| Q)$  is the KL divergence between  $P$  and  $Q$ ,  $D(P_{X|A} \| Q_{X|A} | P_A) = \mathbb{E}_{a \sim P_A} [D(P_{X|A=a} \| Q_{X|A=a})]$ , and  $d_2(p \| q) = p \log \frac{p}{q} + (1-p) \log \frac{1-p}{1-q}$  is the binary KL divergence. To obtain an upper bound, we may take  $p = \frac{p_1}{1+p_1-p_0}$  in (28). Noting that this choice satisfies  $p_0 \leq p \leq p_1$ , we have that

$$\begin{aligned} d_2(p_0 \| p) &= p_0 \log \frac{p_0}{p} + (1-p_0) \log \frac{1-p_0}{1-p} \\ &\leq (1-p_0) \log \frac{1-p_0}{1-p} \\ &= (1-p_0) \log(1+p_1-p_0), \end{aligned} \quad (29)$$

and

$$\begin{aligned} d_2(p_1 \| p) &= p_1 \log \frac{p_1}{p} + (1-p_1) \log \frac{1-p_1}{1-p} \\ &\leq p_1 \log \frac{p_1}{p} \\ &= p_1 \log(1+p_1-p_0), \end{aligned} \quad (30)$$

Substituting (29) and (30) into (28), yields

$$\begin{aligned} & h_2(\alpha p_0 + (1-\alpha)p_1) - \alpha h_2(p_0) - (1-\alpha)h_2(p_1) \\ & \leq \log(1+p_1-p_0) \leq p_1 - p_0. \end{aligned} \quad (31)$$

We will obtain another upper bound using the fact that KL divergence is dominated by  $\chi^2$  divergence. Specifically [35, eq. (5)],  $D(P \| Q) \leq \log(1 + \chi^2(P \| Q))$ . For the binary case, this bound reads

$$d_2(p \| q) \leq \log \left( 1 + \frac{(p-q)^2}{q(1-q)} \right) \leq \frac{(p-q)^2}{q(1-q)}. \quad (32)$$

Now, applying this bound on (28) with the choice  $p = \frac{p_0 + p_1}{2}$ , yields

$$\begin{aligned} & h_2(\alpha p_0 + (1 - \alpha)p_1) - \alpha h_2(p_0) - (1 - \alpha)h_2(p_1) \\ & \leq \frac{(p_1 - p_0)^2}{4 \frac{p_0 + p_1}{2} \left(1 - \frac{p_0 + p_1}{2}\right)} \\ & \leq \frac{(p_1 - p_0)^2}{2 \min\left\{\frac{p_0 + p_1}{2}, 1 - \frac{p_0 + p_1}{2}\right\}} \\ & \leq \frac{(p_1 - p_0)^2}{2 \min\{p_0, 1 - p_1\}}. \end{aligned} \quad (33)$$

Combining (31) and (33) establishes the claim.  $\blacksquare$

Leveraging this result, it is easy to obtain the following ‘‘sphere-packing’’ bound for Bernoulli distributions.

*Corollary 10:* Let  $n \geq 1$  be an integer. Given  $n+1$  numbers  $0 \leq p_0 \leq p_1 \leq \dots \leq p_n \leq 1$ , there must exist an index  $i \in [n]$ , such that for any  $\alpha \in [0, 1]$  we have

$$\begin{aligned} & h_2(\alpha p_{i-1} + (1 - \alpha)p_i) - \alpha h_2(p_{i-1}) - (1 - \alpha)h_2(p_i) \\ & \leq \frac{8}{(n+1)^2}. \end{aligned} \quad (34)$$

*Proof:* Let  $m = \lfloor \frac{n+1}{2} \rfloor$ , and assume  $p_m \leq 1/2$ . We will deal with the case  $p_m > 1/2$  later. Set  $\alpha \in [0, 1]$  and let

$$\begin{aligned} \Delta \triangleq & \min_{i \in [n]} h_2(\alpha p_{i-1} + (1 - \alpha)p_i) \\ & - \alpha h_2(p_{i-1}) - (1 - \alpha)h_2(p_i). \end{aligned} \quad (35)$$

We will show that we must have that  $p_m \geq \frac{\Delta(m+1)^2}{4}$ , and together with  $p_m \leq 1/2$ , this will imply that  $\Delta \leq \frac{8}{(m+1)^2}$ .

By Lemma 9, we have that

$$\Delta \leq \min_{i \in [n]} \min \left\{ p_i - p_{i-1}, \frac{(p_i - p_{i-1})^2}{2 \min\{p_{i-1}, 1 - p_i\}} \right\}. \quad (36)$$

In particular, this implies that for any  $i \in [n]$  we have

$$p_i \geq \max \left\{ p_{i-1} + \Delta, p_{i-1} + \sqrt{2\Delta \min\{p_{i-1}, 1 - p_i\}} \right\}. \quad (37)$$

As  $p_0 \geq 0$ , (37) shows that  $p_1 \geq \Delta$ . Since  $p_i \leq 1/2$  for all  $i \leq m$ , we have that  $\min\{p_{i-1}, 1 - p_i\} = p_{i-1}$  in this range. Thus, for all  $2 \leq i \leq m$ , the second term in (37) dominates the maximum, and

$$p_i \geq p_{i-1} + \sqrt{2\Delta p_{i-1}}. \quad (38)$$

Defining  $\tilde{p}_i = \frac{p_i}{\Delta}$ , we obtain the recursion:

$$\tilde{p}_1 \geq 1 \quad (39)$$

$$\tilde{p}_i \geq \tilde{p}_{i-1} + \sqrt{2\tilde{p}_{i-1}}, \quad 2 \leq i \leq m. \quad (40)$$

It is readily verified by induction that  $\tilde{p}_m \geq \frac{(m+1)^2}{4}$ , which implies that  $p_m \geq \Delta \frac{(m+1)^2}{4}$ , as claimed. Combining this with our assumption that  $p_m \leq 1/2$  and noting that  $m+1 > (n+1)/2$ , we establish (34).

If  $p_m > 1/2$ , we replace each  $p_i$  with  $\bar{p}_i = 1 - p_i$ , such that now  $\bar{p}_m < 1/2$ . Now, reorder (reverse) the probabilities such that  $0 \leq \bar{p}_1 \leq \dots \leq \bar{p}_n$ . We are guaranteed that after

reordering  $\bar{p}_{m'} \leq 1/2$ , where  $m' = n+1 - m = \lceil \frac{n+1}{2} \rceil \geq \lfloor \frac{n+1}{2} \rfloor$ . Since  $h_2(p) = h_2(1-p)$ , we still have

$$\begin{aligned} \Delta = & \min_{i \in [n]} h_2(\alpha \bar{p}_{i-1} + (1 - \alpha)\bar{p}_i) \\ & - \alpha h_2(\bar{p}_{i-1}) - (1 - \alpha)h_2(\bar{p}_i). \end{aligned} \quad (41)$$

Thus, repeating the same argument, we must have that  $\Delta \leq \frac{8}{(m'+1)^2}$ . Combining this with the fact that  $m'+1 > (n+1)/2$ , we establish (34).  $\blacksquare$

## APPENDIX B UPGRADATION FOR BINARY $\mathcal{X}$

We define a procedure that constructs from a joint distribution  $P_{X,Y}$  on  $\mathcal{X} \times \mathcal{Y}$  with  $|\mathcal{X}| = 2$ , a distribution  $P_{X,Z,Y}$  on  $\mathcal{X} \times \mathcal{Z} \times \mathcal{Y}$  with  $|\mathcal{Z}| = |\mathcal{Y}| - 1$ , such that  $\sum_{z \in \mathcal{Z}} P_{X,Z,Y}(x, z, y) = P_{X,Y}(x, y)$ , and  $X - Z - Y$  form a Markov chain in this order.

*Definition 11 (Splitting  $i$ th symbol of  $\mathcal{Y}$ ):* Let  $(X, Y) \sim P_{X,Y}$  be random variables in  $\mathcal{X} \times \mathcal{Y}$ , where  $\mathcal{X} = \{0, 1\}$  and  $\mathcal{Y} = \{1, \dots, |\mathcal{Y}|\}$ . Let  $p_i = \mathbb{P}(X = 0 | Y = i)$  and assume without loss of generality that  $p_1 \leq \dots \leq p_{|\mathcal{Y}|}$ . For any  $i \in \{2, \dots, |\mathcal{Y}| - 1\}$  let  $\alpha_i \in [0, 1]$  satisfy  $\alpha_i p_{i-1} + (1 - \alpha_i)p_{i+1} = p_i$ . The joint distribution  $P_{Y,Z,X}^i = P_Y P_{Z|Y}^i P_{X|Z}^i$  on  $\mathcal{Y} \times (\mathcal{Y} \setminus \{i\}) \times \mathcal{X}$  corresponding to splitting the  $i$ th symbol of  $\mathcal{Y}$  is defined as

$$P_{Z|Y}^i(z|y) = \begin{cases} 1 & y \neq i, z = y \\ \alpha_i & y = i, z = i - 1 \\ 1 - \alpha_i & y = i, z = i + 1 \\ 0 & \text{otherwise,} \end{cases} \quad (42)$$

and  $P_{X|Z}^i(0|z) = p_z$  for all  $z \in \mathcal{Y} \setminus \{i\}$ .

*Proposition 12 (Cost of split):* Under the distribution  $P_{Y,Z,X}^i$ , we have that  $X - Z - Y$  form a Markov chain in this order,  $P_{X,Y}^i = P_{X,Y}$ , and

$$\begin{aligned} I(X; Z) - I(X; Y) = & P_Y(i) \left[ h_2(\alpha_i p_{i-1} + (1 - \alpha_i)p_{i+1}) \right. \\ & \left. - \alpha_i h_2(p_{i-1}) - (1 - \alpha_i)h_2(p_{i+1}) \right]. \end{aligned}$$

*Proof:* Clearly,  $Y - Z - X$  form a Markov chain in this order under  $P_{Y,Z,X}^i$ , and consequently, so do  $X - Z - Y$ . By definition  $P_Y^i = P_Y$ , and therefore, to prove that  $P_{X,Y}^i = P_{X,Y}$ , it suffices to show that  $P_{X|Y}^i = P_{X|Y}$ . To that end, write

$$\begin{aligned} P_{X|Y}^i(0|y) &= \sum_{z \in \mathcal{Y} \setminus \{i\}} P_{X|Z}^i(0|z) P_{Z|Y}^i(z|y) \\ &= \begin{cases} p_y & y \neq i \\ \alpha_i p_{i-1} + (1 - \alpha_i)p_{i+1} & y = i \end{cases} \\ &= P_{X|Y}(0|y), \end{aligned}$$

as  $\alpha_i p_{i-1} + (1 - \alpha_i) p_{i+1} = p_i$  by definition of  $\alpha_i$ . For the mutual information gap, we first have to compute  $P_{Z,X}^i = P_Z^i P_{X|Z}^i$ . To this end, write

$$P_Z^i(z) = \sum_{y \in \mathcal{Y}} P_Y(y) P_{Z|Y}^i(z|y) = \begin{cases} P_Y(z) & z \notin \{i-1, i+1\} \\ P_Y(i-1) + \alpha_i P_Y(i) & z = i-1 \\ P_Y(i+1) + (1 - \alpha_i) P_Y(i) & z = i+1 \end{cases}, \quad (43)$$

and recall that  $P_{X|Z}^i(0|z) = p_z$ . We can now write

$$\begin{aligned} I(X; Z) - I(X; Y) &= H(X|Y) - H(X|Z) \\ &= \sum_{y \in \mathcal{Y}} P_Y(y) h_2(p_y) - \sum_{z \in \mathcal{Y} \setminus \{i\}} P_Z(z) h_2(p_z) \\ &= P_Y(i-1) h_2(p_{i-1}) + P_Y(i) h_2(p_i) + P_Y(i+1) h_2(p_{i+1}) \\ &\quad - (P_Y(i-1) + \alpha_i P_Y(i)) h_2(p_{i-1}) \\ &\quad - (P_Y(i+1) + (1 - \alpha_i) P_Y(i)) h_2(p_{i+1}) \\ &= P_Y(i) h_2(p_i) - \alpha_i P_Y(i) h_2(p_{i-1}) - (1 - \alpha_i) P_Y(i) h_2(p_{i+1}) \\ &= P_Y(i) \left[ h_2(\alpha p_{i-1} + (1 - \alpha_i) p_{i+1}) \right. \\ &\quad \left. - \alpha_i h_2(p_{i-1}) - (1 - \alpha_i) h_2(p_{i+1}) \right], \end{aligned}$$

as claimed.  $\blacksquare$

*Theorem 13:* Let  $(X, Y) \sim P_{X,Y}$  be random variables in  $\mathcal{X} \times \mathcal{Y}$ , where  $\mathcal{X} = \{0, 1\}$  and  $\mathcal{Y} = \{1, \dots, |\mathcal{Y}|\}$ . Let  $p_i = \mathbb{P}(X = 0|Y = i)$  and assume without loss of generality that  $p_1 \leq \dots \leq p_{|\mathcal{Y}|}$ . Then, there exists  $i \in \{2, \dots, |\mathcal{Y}| - 1\}$  such that under the distribution  $P_{Y,Z,X}^i$ , defined in Proposition 12, we have

$$I(X; Z) - I(X; Y) \leq \frac{256}{|\mathcal{Y}|^3}. \quad (44)$$

*Proof:* Without loss of generality we may assume  $|\mathcal{Y}| \geq 8$ , as otherwise the right hand side of (44) is greater than  $\log 2$ , and the statement holds trivially. Let

$$\mathcal{Y}_{\text{small}} \triangleq \left\{ y \in \mathcal{Y} : P_Y(y) \leq \frac{2}{|\mathcal{Y}|} \right\}, \quad (45)$$

which implies  $|\mathcal{Y}_{\text{small}}| \geq \frac{|\mathcal{Y}|}{2}$ , and let  $\mathcal{Y}_{\text{punctured}}$  be the set obtained by removing every other element in  $\mathcal{Y}_{\text{small}}$ , starting from the second. We get that between any two elements of  $\mathcal{Y}_{\text{punctured}}$  lies at least one element of  $\mathcal{Y}_{\text{small}}$ , and that  $|\mathcal{Y}_{\text{punctured}}| \geq \frac{|\mathcal{Y}|}{4}$ . Furthermore, by Corollary 10, we must have two indices  $j, k \in \mathcal{Y}_{\text{punctured}}$ ,  $j > k + 1$ , such that for any  $\alpha \in [0, 1]$

$$\begin{aligned} h_2(\alpha p_k + (1 - \alpha) p_j) - \alpha h_2(p_k) - (1 - \alpha) h_2(p_j) &\leq \frac{8}{\left(\frac{|\mathcal{Y}|}{4}\right)^2} \\ &= \frac{128}{|\mathcal{Y}|^2}. \end{aligned}$$

Thus, we must have some  $i \in \mathcal{Y}_{\text{small}}$  satisfying  $k < i < j$ , for which

$$\begin{aligned} h_2(\alpha_i p_{i-1} + (1 - \alpha_i) p_{i+1}) - \alpha_i h_2(p_{i-1}) - (1 - \alpha_i) h_2(p_{i+1}) \\ \leq \frac{128}{|\mathcal{Y}|^2}, \end{aligned}$$

which follows since<sup>8</sup> for  $0 \leq p \leq p' \leq q' \leq q \leq 1$ ,

$$\begin{aligned} h_2(\alpha p' + (1 - \alpha) q') - \alpha h_2(p') - (1 - \alpha) h_2(q') \\ \leq h_2(\alpha p + (1 - \alpha) q) - \alpha h_2(p) - (1 - \alpha) h_2(q). \end{aligned} \quad (46)$$

By Proposition 12 and (45), for this  $i$  we have that

$$I(X; Z) - I(X; Y) \leq \frac{256}{|\mathcal{Y}|^3},$$

under  $P_{Y,Z,X}^i$ , as claimed.  $\blacksquare$

We are now ready to prove Theorem 1.

*Proof of Theorem 1:* Construct  $Z$  in a greedy fashion as follows. Find an index  $i$  such that  $P_{Y,Z,X}^i$  satisfies (44), and then replace  $Y \leftarrow Z$ , and repeat. Stop when  $|\mathcal{Y}| = L$ . From Theorem 13 we obtain

$$\begin{aligned} I(X; Z) - I(X; Y) &\leq \sum_{\ell=L+1}^{|\mathcal{Y}|} \frac{256}{\ell^3} \\ &< 256 \int_{t=L}^{\infty} \frac{1}{t^3} dt \\ &= \frac{128}{L^2}, \end{aligned}$$

as claimed.  $\blacksquare$

## APPENDIX C

### DEGRADATION FOR BINARY $\mathcal{X}$

For a joint distribution  $P_{X,Y}$  on  $\mathcal{X} \times \mathcal{Y}$ , where  $|\mathcal{X}| = 2$ , we first define a quantizer  $f : \mathcal{Y} \rightarrow \{1, \dots, |\mathcal{Y}| - 1\}$ , with the property that  $I(X; f(Y))$  is close to  $I(X; Y)$ . Then, we apply this quantizer sequentially until the cardinality is reduced to  $L$ .

*Definition 14 (Merging symbols  $i$  and  $j$ ):* Let  $(X, Y) \sim P_{X,Y}$  be random variables in  $\mathcal{X} \times \mathcal{Y}$ , where  $\mathcal{X} = \{0, 1\}$  and  $\mathcal{Y} = \{1, \dots, |\mathcal{Y}|\}$ . The function that merges the symbols  $i \neq j \in \mathcal{Y}$  to the symbol  $y'$ , and does not change the rest of the symbols is defined as  $f_{ij} : \mathcal{Y} \rightarrow \mathcal{Y} \setminus \{i, j\} \cup \{y'\}$ , where  $y' \notin \mathcal{Y}$ . Namely,  $f_{ij}(i) = f_{ij}(j) = y'$  and  $f(y) = y$  for all  $y \notin \{i, j\}$ .

*Proposition 15 (Cost of merge):* Let  $p_i = \mathbb{P}(X = 0|Y = i)$ , and  $\alpha_{ij} = \frac{P_Y(i)}{P_Y(i) + P_Y(j)}$ . Then,

$$\begin{aligned} I(X; Y) - I(X; f_{ij}(Y)) &= (P_Y(i) + P_Y(j)) \\ &\cdot [h_2(\alpha_{ij} p_i + (1 - \alpha_{ij}) p_j) - \alpha_{ij} h_2(p_i) - (1 - \alpha_{ij}) h_2(p_j)]. \end{aligned}$$

*Proof:* Let  $\tilde{Y} = f_{ij}(Y)$ . We clearly have that  $P_{\tilde{Y}}(y) = P_Y(y)$  and  $\mathbb{P}(X = 0|\tilde{Y} = y) = p_i$  for any  $y \in \mathcal{Y} \setminus \{i, j\}$ , while for  $y'$  we have  $P_{\tilde{Y}}(y') = P_Y(i) + P_Y(j)$  and  $\mathbb{P}(X = 0|\tilde{Y} =$

<sup>8</sup>To see this fix  $\alpha$ , and denote the LHS of (46) by  $g(p', q')$ . Recalling that  $0 \leq p' \leq q' \leq 1$ , it suffices to prove that  $g(p', q')$  is non-decreasing if we enlarge  $q'$  or reduce  $p'$ . This is indeed true, and can be seen by considering the partial derivatives of  $g(p', q')$  with respect to  $p'$  and  $q'$ , and noting that  $x/(1-x)$  is increasing in  $x$ , for  $0 \leq x < 1$ .

$y') = \frac{P_Y(i)p_i + P_Y(j)p_j}{P_Y(i) + P_Y(j)} = \alpha_{ij}p_i + (1 - \alpha_{ij})p_j$ . It therefore follows that

$$\begin{aligned} I(X; Y) - I(X; f(Y)) &= H(X|f(Y)) - H(X|Y) \\ &= (P_Y(i) + P_Y(j)) h_2(\alpha_{ij}p_i + (1 - \alpha_{ij})p_j) \\ &+ \sum_{y \in \mathcal{Y} \setminus \{i, j\}} P_Y(y) h_2(p_y) \\ &- P_Y(i) h_2(p_i) - P_Y(j) h_2(p_j) - \sum_{y \in \mathcal{Y} \setminus \{i, j\}} P_Y(y) h_2(p_y) \\ &= (P_Y(i) + P_Y(j)) h_2(\alpha_{ij}p_i + (1 - \alpha_{ij})p_j) \\ &- P_Y(i) h_2(p_i) - P_Y(j) h_2(p_j), \end{aligned}$$

and the result follows by definition of  $\alpha_{ij}$ .  $\blacksquare$

The following theorem shows that if  $|\mathcal{Y}|$  is large, we can always find two symbols such that merging them would not significantly decrease the mutual information.

*Remark 16:* The merging operation is in fact a quantizer of  $Y$  to  $|\mathcal{Y}| - 1$  levels. It is known [21], [36], [37] that an optimal quantizer  $f: \mathcal{Y} \rightarrow [M]$ , in terms of maximizing  $I(X; f(Y))$ , can be associated with a partition of the interval  $[0, 1]$  to  $M$  disjoint intervals  $\mathcal{I}_1, \dots, \mathcal{I}_M$ , such that  $f(y) = m$  iff  $\mathbb{P}(X = 0|Y = y) \in \mathcal{I}_m$ . Thus, if we relabel  $\mathcal{Y} = \{1, \dots, |\mathcal{Y}|\}$  such that  $p_1 \leq \dots \leq p_{|\mathcal{Y}|}$ , we have that the symbols with the smallest cost of merge are adjacent. It therefore suffices to restrict the search to  $i \in \{1, \dots, |\mathcal{Y}| - 1\}$  and  $j = i + 1$ .

*Theorem 17:* Let  $(X, Y) \sim P_{X, Y}$  be random variables in  $\mathcal{X} \times \mathcal{Y}$ , where  $\mathcal{X} = \{0, 1\}$ ,  $\mathcal{Y} = \{1, \dots, |\mathcal{Y}|\}$ , and let  $f_{ij}$  be as defined in Proposition 15. Then, there exists  $i \neq j \in \mathcal{Y}$  such that

$$I(X; Y) - I(X; f_{ij}(Y)) \leq \frac{128}{|\mathcal{Y}|^3}. \quad (47)$$

*Proof:* Without loss of generality we may assume  $|\mathcal{Y}| \geq 4$ , as otherwise the right hand side of (47) is greater than  $\log 2$  and the statement holds trivially. Let  $\mathcal{Y}_{\text{small}}$  be as in (45), and recall that  $|\mathcal{Y}_{\text{small}}| \geq \frac{|\mathcal{Y}|}{2}$ . By Corollary 10, there must exist  $i \neq j \in \mathcal{Y}_{\text{small}} \subset \mathcal{Y}$  for which

$$\begin{aligned} h_2(\alpha_{ij}p_i + (1 - \alpha_{ij})p_j) - \alpha_{ij}h_2(p_i) - (1 - \alpha_{ij})h_2(p_j) \\ \leq \frac{8}{|\mathcal{Y}_{\text{small}}|^2} \leq \frac{32}{|\mathcal{Y}|^2}. \end{aligned} \quad (48)$$

Furthermore, as  $i, j \in \mathcal{Y}_{\text{small}}$ , we have that

$$P_Y(i) + P_Y(j) \leq \frac{4}{|\mathcal{Y}|}. \quad (49)$$

The claim now immediately follows from substituting (48) and (49) into Proposition 15.  $\blacksquare$

*Theorem 18:* Let  $(X, Y) \sim P_{X, Y}$  be random variables in  $\mathcal{X} \times \mathcal{Y}$ , where  $|\mathcal{X}| = 2$  and  $\mathcal{Y}$  is discrete. For any integer  $L$ , the greedy-merge algorithm finds a quantizer  $f: \mathcal{Y} \rightarrow \{1, \dots, L\}$  such that

$$I(X; Y) - I(X; f(Y)) \leq \frac{64}{L^2}. \quad (50)$$

*Proof:* Construct  $f$  in a greedy fashion as follows. Merge the two symbols of  $\mathcal{Y}$  for which the loss in mutual information

due to merging is smallest, and repeat this until  $|\mathcal{Y}| = L$ . From Theorem 17 we obtain

$$\begin{aligned} I(X; Y) - I(X; f(Y)) &\leq \sum_{\ell=L+1}^{|\mathcal{Y}|} \frac{128}{\ell^3} \\ &< 128 \int_{t=L}^{\infty} \frac{1}{t^3} dt \\ &= \frac{64}{L^2}, \end{aligned} \quad (51)$$

as claimed.  $\blacksquare$

#### APPENDIX D DEGRADATION FOR GENERAL $\mathcal{X}$

Repeating the proof of [28, Proposition 7], with the improved performance guarantees of the greedy merge algorithm for binary  $\mathcal{X}$ , as stated in Theorem 18, yields the bound (27) from Theorem 8. For completeness, we bring the proof below.

*Proof of (27) from Theorem 8:* The case  $q = 2$  is covered by Theorem 18.

Now let  $q > 2$ , and without loss of generality assume  $\mathcal{X} = \{1, 2, \dots, q\}$ . Define  $X_i \triangleq \mathbf{1}_{\{X=i\}}$ , for  $i = 1, 2, \dots, q - 1$ . Then,

$$\begin{aligned} I(X; Y) &= I(X_1, \dots, X_{q-1}; Y) \\ &= \sum_{i=1}^{q-1} I(X_i; Y | X_1^{i-1} = 0_1^{i-1}) \mathbb{P}(X_1^{i-1} = 0_1^{i-1}) \end{aligned} \quad (52)$$

where  $X_1^{i-1} = 0_1^{i-1}$  denotes the event  $X_1 = \dots = X_{i-1} = 0$ .

Let  $f(y)$  be an  $M$ -level quantizer, with  $M \leq L$ , of the form  $f(y) = (f_1(y), \dots, f_{q-1}(y))$ . Then,

$$\begin{aligned} I(X; f(Y)) &= \sum_{i=1}^{q-1} I(X_i; f(Y) | X_1^{i-1} = 0_1^{i-1}) \mathbb{P}(X_1^{i-1} = 0_1^{i-1}) \\ &\geq \sum_{i=1}^{q-1} I(X_i; f_i(Y) | X_1^{i-1} = 0_1^{i-1}) \mathbb{P}(X_1^{i-1} = 0_1^{i-1}). \end{aligned} \quad (53)$$

Thus, combining (52) and (53), gives

$$\begin{aligned} I(X; Y) - I(X; f(Y)) &\leq \sum_{i=1}^{q-1} (I(X_i; Y | X_1^{i-1} = 0_1^{i-1}) \\ &- I(X_i; f_i(Y) | X_1^{i-1} = 0_1^{i-1})) \mathbb{P}(X_1^{i-1} = 0_1^{i-1}). \end{aligned} \quad (54)$$

It follows from Theorem 18 that by taking the support sizes as  $|f_i(y)| = \Lambda$ , where  $\Lambda = \lfloor L^{1/(q-1)} \rfloor$ , for all  $1 \leq i \leq q - 1$ , that the quantizers  $f_1(y), \dots, f_{q-1}(y)$  returned by the greedy-merge algorithm satisfy

$$I(X_i; Y | X_1^{i-1} = 0_1^{i-1}) - I(X_i; f_i(Y) | X_1^{i-1} = 0_1^{i-1}) \leq \frac{64}{\Lambda^2}.$$

Consequently, with this choice, we obtain

$$\begin{aligned} I(X; Y) - I(X; f(Y)) &\leq 64 \cdot \Lambda^{-2} \sum_{i=1}^{q-1} \mathbb{P}(X_1^{i-1} = 0) \\ &\leq 64(q-1) \cdot \left[ L^{1/(q-1)} \right]^{-2}, \end{aligned} \quad (55)$$

as desired.  $\blacksquare$

## REFERENCES

- [1] O. Ordentlich and I. Tal, "An upgrading algorithm with optimal power law," in *International Zurich Seminar on Information and Communication (IZS 2020)*. ETH Zurich, 2020, p. 46.
- [2] E. Arkan, "Channel polarization: A method for constructing capacity-achieving codes for symmetric binary-input memoryless channels," *IEEE Trans. Inform. Theory*, vol. 55, no. 7, pp. 3051–3073, July 2009.
- [3] R. Mori and T. Tanaka, "Performance and construction of polar codes on symmetric binary-input memoryless channels," in *Proc. IEEE Int'l Symp. Inform. Theory (ISIT'2009)*, Seoul, South Korea, 2009, pp. 1496–1500.
- [4] S. B. Korada, "Polar codes for channel and source coding," Ph.D. dissertation, Ecole Polytechnique Fédérale de Lausanne, 2009.
- [5] I. Tal and A. Vardy, "How to construct polar codes," *IEEE Trans. Inform. Theory*, vol. 59, no. 10, pp. 6562–6582, October 2013.
- [6] S. B. Korada and R. Urbanke, "Polar codes are optimal for lossy source coding," *IEEE Trans. Inform. Theory*, vol. 56, no. 4, pp. 1751–1768, April 2010.
- [7] J. Honda and H. Yamamoto, "Polar coding without alphabet extension for asymmetric channels," *IEEE Trans. Inform. Theory*, vol. 59, no. 12, pp. 7829–7838, December 2012.
- [8] E. Hof and S. Shamai, "Secrecy-achieving polar-coding for binary-input memoryless symmetric wire-tap channels," arXiv:1005.2759v2, 2010.
- [9] H. Mahdavi and A. Vardy, "Achieving the secrecy capacity of wiretap channels using polar codes," *IEEE Trans. Inform. Theory*, vol. 57, pp. 6428–6443, 2011.
- [10] M. Andersson, V. Rathi, R. Thobaben, J. Kliewer, and M. Skoglund, "Nested polar codes for wiretap and relay channels," *IEEE Commun. Lett.*, vol. 14, pp. 752–754, 2010.
- [11] O. O. Koyluoglu and H. El Gamal, "Polar coding for secure transmission and key agreement," *IEEE Transactions on Information Forensics and Security*, vol. 7, no. 5, pp. 1472–1483, October 2012.
- [12] J. M. Renes, R. Renner, and D. Sutter, "Efficient one-way secret-key agreement and private channel coding via polarization," in *International Conference on the Theory and Application of Cryptology and Information Security*. Springer, 2013, pp. 194–213.
- [13] R. Wang, J. Honda, H. Yamamoto, R. Liu, and Y. Hou, "Construction of polar codes for channels with memory," in *Proc. IEEE Inform. Theory Workshop (ITW'2015)*, Jeju Island, Korea, 2015, pp. 187–191.
- [14] E. Şaşoğlu and I. Tal, "Polar coding for processes with memory," *IEEE Trans. Inform. Theory*, vol. 65, no. 4, pp. 1994–2003, April 2019.
- [15] B. Shuval and I. Tal, "Fast polarization for processes with memory," *IEEE Trans. Inform. Theory*, vol. 65, no. 4, pp. 2004–2020, April 2019.
- [16] —, "Universal polarization for processes with memory," arXiv:1811.05727v1, 2018.
- [17] I. Tal, H. D. Pfister, A. Fazeli, and A. Vardy, "Polar codes for the deletion channel: weak and strong polarization," arXiv:1904.13385v1, 2019.
- [18] R. Pedarsani, S. H. Hassani, I. Tal, and E. Telatar, "On the construction of polar codes," in *Proc. IEEE Int'l Symp. Inform. Theory (ISIT'2011)*, Saint Petersburg, Russia, 2011, pp. 11–15.
- [19] I. Tal and A. Vardy, "Channel upgrading for semantically-secure encryption on wiretap channels," in *Proc. IEEE Int'l Symp. Inform. Theory (ISIT'2013)*, Istanbul, Turkey, 2013.
- [20] T. Koch and A. Lapidoth, "At low SNR, asymmetric quantizers are better," *IEEE Trans. Inform. Theory*, vol. 59, no. 9, pp. 5421–5445, September 2013.
- [21] B. M. Kurkoski and H. Yagi, "Quantization of binary-input discrete memoryless channels," *IEEE Trans. Inform. Theory*, vol. 60, no. 8, pp. 4544–4552, August 2014.
- [22] G. Alirezaei and R. Mathar, "Optimal one-bit quantizers are asymmetric for additive uniform noise," in *Information Theory and Applications Workshop (ITA'2017)*, 2017.
- [23] I. Tal, A. Sharov, and A. Vardy, "Constructing polar codes for non-binary alphabets and MACs," in *Proc. IEEE Int'l Symp. Inform. Theory (ISIT'2012)*, Cambridge, Massachusetts, 2012, pp. 2132–2136.
- [24] U. Pereg and I. Tal, "Channel upgrading for non-binary input alphabets and MACs," *IEEE Trans. Inform. Theory*, vol. 63, no. 3, pp. 1410–1424, March 2017.
- [25] T. C. Gulcu, M. Ye, and A. Barg, "Construction of polar codes for arbitrary discrete memoryless channels," *IEEE Trans. Inform. Theory*, vol. 64, pp. 309–321, January 2017.
- [26] I. Tal, "On the construction of polar codes for channels with moderate input alphabet sizes," *IEEE Trans. Inform. Theory*, vol. 63, no. 3, pp. 1501–1509, March 2017.
- [27] A. Kartowsky and I. Tal, "Greedy-merge degrading has optimal power-law," *IEEE Trans. Inform. Theory*, vol. 65, no. 2, pp. 917–934, February 2019.
- [28] A. Bhatt, B. Nazer, O. Ordentlich, and Y. Polyanskiy, "Information-distilling quantizers," *IEEE Trans. Inform. Theory*, vol. 67, no. 4, pp. 2472–2487, 2021.
- [29] A. Zhang and B. M. Kurkoski, "Low-complexity quantization of discrete memoryless channels," *Proc. IEEE Int'l Symp. Inform. Theory Appl. (ISITA'2016)*, 2016.
- [30] K. I. Iwata and S. Y. Ozawa, "Quantizer design for outputs of binary-input discrete memoryless channels using SMAWK algorithm," in *Proc. IEEE Int'l Symp. Inform. Theory (ISIT'2014)*, June 2014, pp. 191–195.
- [31] Y. Sakai and K. I. Iwata, "Optimal quantization of B-DMCs maximizing  $\alpha$ -mutual information with monge property," in *Proc. IEEE Int'l Symp. Inform. Theory (ISIT'2017)*, June 2017.
- [32] N. Slonim and N. Tishby, "Agglomerative information bottleneck," in *NIPS*, vol. 4, 1999.
- [33] I. Tal, "polarcub," <https://github.com/idotal/polarcub>, 2021.
- [34] I. Csiszar and J. Körner, *Information theory: coding theorems for discrete memoryless systems*. Academic Press, 1981.
- [35] I. Sason and S. Verdú, " $f$ -divergence inequalities," *IEEE Transactions on Information Theory*, vol. 62, no. 11, pp. 5973–6006, Nov 2016.
- [36] D. Burshtein, V. Della Pietra, D. Kanevsky, and A. Nadas, "Minimum impurity partitions," *The Annals of Statistics*, vol. 20, no. 3, pp. 1637–1646, 1992.
- [37] X. Wu, P. Chou, and X. Xue, "Minimum conditional entropy context quantization," in *2000 IEEE International Symposium on Information Theory*, 2000, p. 43.